Scalar products in $GL(3)$-based models with trigonometric R-matrix. Determinant representation

N. A. Slavnov[1]

Steklov Mathematical Institute, Moscow, Russia

Abstract

We study quantum integrable $GL(3)$-based models with a trigonometric R-matrix solvable by the nested algebraic Bethe ansatz. We derive a determinant representation for a special case of scalar products of Bethe vectors. This representation allows one to find a determinant formula for the form factor of one of the monodromy matrix entries. We also point out essential difference between form factors in the models with the trigonometric R-matrix and their analogs in $GL(3)$-invariant models.

1 Introduction

The algebraic Bethe ansatz [1–4] allows one to obtain the spectra of quantum Hamiltonians of many models of physical interest. The calculation of correlation functions and form factors also can be performed in the framework of this method. The last problem in many cases can be reduced to the calculation of scalar products of Bethe vectors. A systematic study of scalar products in quantum integrable $GL(3)$-based models with a trigonometric R-matrix [5, 6] was initiated in [7, 8].

Recently, the form factors of the monodromy matrix entries in models with $GL(3)$-invariant R-matrix were studied in works [9–13]. These form factors can be used for calculating correlation functions in a wide class of Bethe ansatz solvable models. It was shown in [13] that all form factors of the monodromy matrix entries are related to each other. Actually, it is enough to calculate only one of them. All the others can be obtained from this initial one within the special limits of Bethe parameters. This method was developed in [13], where it was called *a zero modes method*. In fact, this method allows one to obtain all the form factors from a scalar product of Bethe vectors of a special type.

A question arises about a generalization of these results to the models with trigonometric (or q-deformed) R-matrix. The trigonometric quantum R-matrix for the $GL(3)$-based models

[1] nslavnov@mi.ras.ru

has the following form:

$$
\begin{aligned}
R(u,v) \; &= \; f(u,v) \sum_{1\le i\le 3} \mathsf{E}_{ii}\otimes \mathsf{E}_{ii} \; + \sum_{1\le i<j\le 3} (\mathsf{E}_{ii}\otimes \mathsf{E}_{jj}+\mathsf{E}_{jj}\otimes \mathsf{E}_{ii}) \\
&\qquad + \sum_{1\le i<j\le 3} (u\, g(u,v)\mathsf{E}_{ij}\otimes \mathsf{E}_{ji}+v\, g(u,v)\mathsf{E}_{ji}\otimes \mathsf{E}_{ij}),
\end{aligned}
\tag{1.1}
$$

where the rational functions $f(u,v)$ and $g(u,v)$ are

$$
f(u,v)=\frac{qu-q^{-1}v}{u-v}, \quad g(u,v)=\frac{q-q^{-1}}{u-v}\,,
\tag{1.2}
$$

and q is a complex number (a deformation parameter). The matrices $(\mathsf{E}_{ij})_{lk}=\delta_{il}\delta_{jk}$, $i,j,l,k=1,2,3$ of the size 3×3 have a unit in the intersection of ith row and jth column and zero matrix elements elsewhere.

For the $GL(2)$-based models q-deformation of the R-matrix does not lead to serious problems in calculating the scalar products of Bethe vectors. In particular, there exists a universal determinant representation for the scalar product of an arbitrary Bethe vector with an eigenstate of the transfer matrix [14]. Being written in terms of the functions (1.2) this representation is valid both for the models with $GL(2)$-invariant R-matrix and for the models with the q-deformed symmetry. In the first case one should only take the scaling limit in (1.2): $u=1+\epsilon u'$, $v=1+\epsilon v'$, $q=1+\epsilon c/2$, $\epsilon\to 0$.

One could expect that this similarity is preserved in the models described by the q-deformation of higher rank algebras. This expectation was partly confirmed in the recent works devoted to the analysis of scalar products in the $GL(3)$-based models with the trigonometric R-matrix [7, 8]. In particular, there was obtained a q-deformed generalization of the *sum formula* [15] for the scalar product of generic Bethe vectors. In this representation the scalar products are given as sums over partitions of the sets of Bethe parameters. It is known, however, that this type of representations are not convenient either for analytical calculations of correlation functions or for their numeric analysis. It would be desirable to reduce the sums over partitions of Bethe parameters to determinant formulas similar to the ones obtained in [9–13] for the models with $GL(3)$-invariant R-matrix. The main goal of this paper is to obtain these representations.

To our great surprise it turned out that the q-deformation of $GL(3)$-invariance leads to significant changes at the level of determinant representations for form factors. In the present paper we obtain this representation for form factor of only one entry of the monodromy matrix. Furthermore, we show that obtaining determinant representations for other form factors is impossible, at least in the framework of the scheme considered below. Instead, one can derive determinant formulas for so called *twisted form factors*, where one of the vectors is an eigenstate of a twisted transfer matrix (see section 4) with a special twist parameter.

The paper is organized as follows. In section 2 we describe the model under consideration and introduce a necessary notation. Section 3 is devoted to the description of Bethe vectors and their scalar products. In section 4 we introduce a notion of a twisted transfer matrix that we use for calculating form factors. In section 5 we present the main result of the paper: a determinant representation for a special type of the scalar product of Bethe vectors in the $GL(3)$-based models with the trigonometric R-matrix. The remaining part of the paper is devoted to the

proof of these formulas. In section 6 we describe some properties of the Izergin determinant and formulate several lemmas used in the next sections. In sections 7 and 8 we derive a determinant formula for the scalar product of twisted on-shell and usual on-shell Bethe vectors. Section 9 is devoted to the conclusive discussions. Appendices A and B contain the proofs of the lemmas formulated in section 6.

2 Notation

In this paper we consider quantum integrable models with a 3×3 monodromy matrix $T(u)$. It satisfies standard commutation relation (RTT-relation):

$$R(u,v) \cdot (T(u) \otimes \mathbf{1}) \cdot (\mathbf{1} \otimes T(v)) = (\mathbf{1} \otimes T(v)) \cdot (T(u) \otimes \mathbf{1}) \cdot R(u,v)\,. \tag{2.1}$$

The entries $T_{ij}(u)$ of the monodromy matrix act in a quantum space V and form the quadratic algebra with commutation relations given by (2.1). We assume that the vector space V possesses a highest weight vector $|0\rangle \in V$ such that:

$$T_{ij}(u)|0\rangle = 0, \quad i > j, \quad T_{ii}(u)|0\rangle = \lambda_i(u)|0\rangle\,, \qquad \lambda_i(u) \in \mathbb{C}[[u, u^{-1}]]\,. \tag{2.2}$$

We also assume that the operators $T_{ij}(u)$ act in a dual space V^* with a vector $\langle 0| \in V^*$ such that:

$$\langle 0|T_{ij}(u) = 0, \quad i < j, \quad \langle 0|T_{ii}(u) = \lambda_i(u)\langle 0|\,, \tag{2.3}$$

and $\lambda_i(u)$ are the same as in (2.2). Actually, it is always possible to normalize the monodromy matrix $T(u) \to \lambda_2^{-1}(u)T(u)$ so as to deal only with the ratios

$$r_1(u) = \frac{\lambda_1(u)}{\lambda_2(u)}, \qquad r_3(u) = \frac{\lambda_3(u)}{\lambda_2(u)}. \tag{2.4}$$

Below we assume that $\lambda_2(u) = 1$.

Apart from the functions $g(x,y)$ and $f(x,y)$ (1.2) we also use two other auxiliary functions

$$h(x,y) = \frac{f(x,y)}{g(x,y)} = \frac{qx - q^{-1}y}{q - q^{-1}}, \qquad t(x,y) = \frac{g(x,y)}{h(x,y)} = \frac{(q-q^{-1})^2}{(x-y)(qx - q^{-1}y)}. \tag{2.5}$$

The following obvious properties of the functions introduced above are useful:

$$\begin{aligned} &h(xq^{-2},y) = q^{-1}g^{-1}(x,y), \qquad &&g(x,yq^{-2}) = qh^{-1}(x,y), \qquad &&g(xq^{-2},x) = -qx^{-1},\\ &t(xq^{-2},y) = q^2 t(y,x), &&t(x,yq^2) = q^{-2}t(y,x), &&f(xq^{-2},y) = f^{-1}(y,x). \end{aligned} \tag{2.6}$$

Below we permanently deal with sets of variables and their partitions into subsets. We denote sets of variables by bar: $\bar u$, $\bar v$ and so on,

$$\bar u = \{u_1, \dots, u_a\}\,, \quad \bar v = \{v_1, \dots, v_b\}\ . \tag{2.7}$$

If necessary, the cardinalities of the sets are described in special comments after the formulas. The individual elements of the sets are denoted by subscripts: u_j, v_k etc. Special notation $\bar u_j$,

$\bar{v}_k$ and similar ones are used for subsets with one element omitted: $\bar{u}_j = \bar{u} \setminus u_j$, $\bar{v}_k = \bar{v} \setminus v_k$ and so on.

If a set of variables is multiplied by a number $\alpha\bar{u}$ (in particular, $\bar{u}q^{\pm 2}$), then it means that all the elements of the set are multiplied by this number:

$$\alpha\bar{u} = \{\alpha u_1, \dots, \alpha u_a\}, \qquad \bar{v}q^{\pm 2} = \left\{v_1 q^{\pm 2}, \dots, v_b q^{\pm 2}\right\}. \tag{2.8}$$

If the order of the elements in a set is essential, then we assume that the elements are ordered in such a way that the sequence of their subscripts is strictly increasing. We call this ordering a natural order.

A union of sets is denoted by braces, for example, $\{\bar{w}, \bar{u}\} = \bar{\eta}$. Partitions of sets into disjoint subsets are denoted by the symbol $\Rightarrow$, and the subsets are numerated by roman numbers. For example, notation $\bar{u} \Rightarrow \{\bar{u}_{\rm I}, \bar{u}_{\rm II}\}$ means that the set $\bar{u}$ is divided into two subsets $\bar{u}_{\rm I}$ and $\bar{u}_{\rm II}$, such that $\bar{u}_{\rm I} \cap \bar{u}_{\rm II} = \emptyset$ and $\{\bar{u}_{\rm I}, \bar{u}_{\rm II}\} = \bar{u}$. Similarly, notation $\bar{\eta} \Rightarrow \{\bar{\eta}_{\rm i}, \bar{\eta}_{\rm ii}, \bar{\eta}_{\rm iii}\}$ means that the set $\bar{\eta}$ is divided into three subsets with pair-wise empty intersections and $\{\bar{\eta}_{\rm i}, \bar{\eta}_{\rm ii}, \bar{\eta}_{\rm iii}\} = \bar{\eta}$.

Just like in [7] we use a shorthand notation for products with respect to sets of variables:

$$T_{ij}(\bar{w}) = \prod_{w_k \in \bar{w}} T_{ij}(w_k); \quad r_1(\bar{v}_k) = \prod_{\substack{v_j \in \bar{v} \\ v_j \ne v_k}} r_1(v_j); \quad r_3(\bar{v}_{\rm I}) = \prod_{v_j \in \bar{v}_{\rm I}} r_3(v_j). \tag{2.9}$$

That is, if the operator T_{ij} or the functions r_k depend on a set of variables, this means that one should take the product of the operators or the scalar functions with respect to the corresponding set. The same convention will be used for the products of functions $f(u,v)$, $h(u,v)$, and $t(u,v)$:

$$h(u, \bar{w}) = \prod_{w_j \in \bar{w}} h(u, w_j); \quad t(\bar{v}_k, u) = \prod_{\substack{v_j \in \bar{v} \\ v_j \ne v_k}} t(v_j, u); \quad f(\bar{u}_{\rm I}, \bar{v}_{\rm II}) = \prod_{u_j \in \bar{u}_{\rm I}} \prod_{v_k \in \bar{v}_{\rm II}} f(u_j, v_k). \tag{2.10}$$

In addition to the standard shorthand notations we will also use

$$\mathfrak{p}(\bar{x}) = \prod_{x_k \in \bar{x}} x_k\,. \tag{2.11}$$

3 Bethe vectors

Bethe vectors are special polynomials in the operators $T_{jk}(w)$ with $j < k$ applied to the vector $|0\rangle$. Similarly, dual Bethe vectors are special polynomials in the operators $T_{jk}(w)$ with $j > k$ applied to $\langle 0|$. Their explicit form in the $GL(N)$-based models with the trigonometric R-matrix was found in [16]. We denote Bethe vectors and their dual ones respectively by $\mathbb{B}^{a,b}(\bar{u};\bar{v})$ and $\mathbb{C}^{a,b}(\bar{u};\bar{v})$, stressing that they depend on two sets of variables $\bar{u}$ and $\bar{v}$. These variables are called Bethe parameters. The superscripts a and b show the cardinalities of the sets $\bar{u}$ and $\bar{v}$: $\#\bar{u} = a$, $\#\bar{v} = b$.

A Bethe vector[2] becomes an eigenvector of the transfer matrix $\mathrm{tr}\, T(u)$ if the Bethe parameters satisfy a system of Bethe equations

$$r_1(u_j) = \frac{f(u_j, \bar{u}_j)}{f(\bar{u}_j, u_j)} f(\bar{v}, u_j), \qquad r_3(v_j) = \frac{f(\bar{v}_j, v_j)}{f(v_j, \bar{v}_j)} f(v_j, \bar{u}), \tag{3.1}$$

[2]For simplicity below we do not distinguish between Bethe vectors and dual ones, if it does not cause a misunderstanding.

and we recall that $\bar{u}_j = \bar{u} \setminus u_j$, $\bar{v}_j = \bar{v} \setminus v_j$. In this case we call the corresponding vector *on-shell* Bethe vector. Observe that if $\bar{u}$ and $\bar{v}$ satisfy (3.1), then for arbitrary partitions $\bar{u} \Rightarrow \{\bar{u}_{\rm I}, \bar{u}_{\rm II}\}$ and $\bar{v} \Rightarrow \{\bar{v}_{\rm I}, \bar{v}_{\rm II}\}$

$$r_1(\bar{u}_{\rm I}) = \frac{f(\bar{u}_{\rm I}, \bar{u}_{\rm II})}{f(\bar{u}_{\rm II}, \bar{u}_{\rm I})} f(\bar{v}, \bar{u}_{\rm I}), \qquad r_3(\bar{v}_{\rm I}) = \frac{f(\bar{v}_{\rm II}, \bar{v}_{\rm I})}{f(\bar{v}_{\rm I}, \bar{v}_{\rm II})} f(\bar{v}_{\rm I}, \bar{u}). \tag{3.2}$$

To obtain (3.2) it is enough to take the products in (3.1) with respect to the subsets $\bar{u}_{\rm I}$ and $\bar{v}_{\rm I}$.

If the system of Bethe equations (3.1) holds, then

$$\operatorname{tr} T(w) \mathbb{B}^{a,b}(\bar{u}; \bar{v}) = \tau(w|\bar{u}, \bar{v})\, \mathbb{B}^{a,b}(\bar{u}; \bar{v}), \qquad \mathbb{C}^{a,b}(\bar{u}; \bar{v}) \operatorname{tr} T(w) = \tau(w|\bar{u}, \bar{v})\, \mathbb{C}^{a,b}(\bar{u}; \bar{v}), \tag{3.3}$$

where

$$\tau(w|\bar{u}, \bar{v}) = r_1(w) f(\bar{u}, w) + f(w, \bar{u}) f(\bar{v}, w) + r_3(w) f(w, \bar{v}). \tag{3.4}$$

Scalar products of Bethe vectors are defined as follows:

$$S_{a,b} \equiv S_{a,b}(\bar{u}^C, \bar{v}^C; \bar{u}^B, \bar{v}^B) = \mathbb{C}^{a,b}(\bar{u}^C; \bar{v}^C) \mathbb{B}^{a,b}(\bar{u}^B; \bar{v}^B). \tag{3.5}$$

Here the vectors $\mathbb{C}^{a,b}(\bar{u}^C; \bar{v}^C)$ and $\mathbb{B}^{a,b}(\bar{u}^B; \bar{v}^B)$ supposed to be generic Bethe vectors, i.e. their Bethe parameters are generic complex numbers. We have added the superscripts C and B to the sets $\bar{u}$, $\bar{v}$ in order to stress that the vectors $\mathbb{C}^{a,b}(\bar{u}^C; \bar{v}^C)$ and $\mathbb{B}^{a,b}(\bar{u}^B; \bar{v}^B)$ may depend on different sets of parameters. In other words, unless explicitly specified, the variables $\{\bar{u}^B, \bar{v}^B\}$ in $\mathbb{B}^{a,b}$ and $\{\bar{u}^C, \bar{v}^C\}$ in $\mathbb{C}^{a,b}$ are not supposed to be equal.

The main goal of this paper is to find determinant representations for form factors of the diagonal entries $T_{ii}(z)$ of the monodromy matrix. We denote them by $\mathcal{F}^{(i,i)}_{a,b}(z)$ and define as

$$\mathcal{F}^{(i,i)}_{a,b}(z) \equiv \mathcal{F}^{(i,i)}_{a,b}(z|\bar{u}^C, \bar{v}^C; \bar{u}^B, \bar{v}^B) = \mathbb{C}^{a,b}(\bar{u}^C; \bar{v}^C) T_{ii}(z) \mathbb{B}^{a,b}(\bar{u}^B; \bar{v}^B), \tag{3.6}$$

where both $\mathbb{C}^{a,b}(\bar{u}^C; \bar{v}^C)$ and $\mathbb{B}^{a,b}(\bar{u}^B; \bar{v}^B)$ are on-shell Bethe vectors. The parameter z is an arbitrary complex number.

4 Twisted transfer matrix

One can easily relate the form factors of the diagonal matrix elements $T_{ii}(z)$ with scalar products of a special type [9, 10, 17–19]. For this we introduce the notion of a twisted monodromy matrix and a twisted transfer matrix.

Let $\bar{\kappa}$ be a set of three complex numbers (twist parameters) $\bar{\kappa} = \{\kappa_1, \kappa_2, \kappa_3\}$. We define the twisted monodromy matrix as $T_{\bar{\kappa}} = \hat{\kappa} T$, where $\hat{\kappa} = \operatorname{diag}(\kappa_1, \kappa_2, \kappa_3)$. It is easy to check that the tensor square of $\hat{\kappa}$ commutes with the R-matrix: $[\hat{\kappa}_1 \hat{\kappa}_2, R_{12}(u, v)] = 0$, and therefore the twisted monodromy matrix satisfies the algebra (2.1)

$$R(u, v) \cdot (T_{\bar{\kappa}}(u) \otimes \mathbf{1}) \cdot (\mathbf{1} \otimes T_{\bar{\kappa}}(v)) = (\mathbf{1} \otimes T_{\bar{\kappa}}(v)) \cdot (T_{\bar{\kappa}}(u) \otimes \mathbf{1}) \cdot R(u, v)\,. \tag{4.1}$$

Thus, the eigenvectors of the twisted transfer matrix $\operatorname{tr} T_{\bar{\kappa}}(w)$ can be found within the framework of the standard scheme. We call them twisted (dual) on-shell Bethe vectors. Like the standard

on-shell vectors, they can be parameterized by sets of complex parameters satisfying twisted Bethe equations

$$r_1(u_j) = \frac{\kappa_2}{\kappa_1}\frac{f(u_j,\bar u_j)}{f(\bar u_j,u_j)}f(\bar v,u_j), \qquad r_3(v_j) = \frac{\kappa_2}{\kappa_3}\frac{f(\bar v_j,v_j)}{f(v_j,\bar v_j)}f(v_j,\bar u). \tag{4.2}$$

Similarly to (3.2) one can find for arbitrary partitions $\bar u \Rightarrow \{\bar u_{\rm I},\bar u_{\rm II}\}$ and $\bar v \Rightarrow \{\bar v_{\rm I},\bar v_{\rm II}\}$

$$r_1(\bar u_{\rm I}) = \left(\frac{\kappa_2}{\kappa_1}\right)^{k_{\rm I}}\frac{f(\bar u_{\rm I},\bar u_{\rm II})}{f(\bar u_{\rm II},\bar u_{\rm I})}f(\bar v,\bar u_{\rm I}), \qquad r_3(\bar v_{\rm I}) = \left(\frac{\kappa_2}{\kappa_3}\right)^{n_{\rm I}}\frac{f(\bar v_{\rm II},\bar v_{\rm I})}{f(\bar v_{\rm I},\bar v_{\rm II})}f(\bar v_{\rm I},\bar u), \tag{4.3}$$

where $k_{\rm I} = \#\bar u_{\rm I}$ and $n_{\rm I} = \#\bar v_{\rm I}$.

In this paper we will also consider twisted from factors. They are still defined by (3.6), but one of vectors (for instance, $\mathbb{C}^{a,b}(\bar u^C;\bar v^C)$) is a twisted on-shell Bethe vector. Calculating the twisted and usual form factors of the matrix elements $T_{ii}(z)$ can be reduced to the scalar products via the following method.

Let $\operatorname{tr} T_{\bar\kappa}(z)$ be the twisted transfer matrix and $\operatorname{tr} T(z)$ be the standard transfer matrix. Consider

$$Q_{\bar\kappa}(z) = \mathbb{C}^{a,b}(\bar u^C;\bar v^C)\big(\operatorname{tr} T_{\bar\kappa}(z) - \operatorname{tr} T(z)\big)\mathbb{B}^{a,b}(\bar u^B;\bar v^B), \tag{4.4}$$

where $\mathbb{C}^{a,b}(\bar u^C;\bar v^C)$ and $\mathbb{B}^{a,b}(\bar u^B;\bar v^B)$ are twisted and standard on-shell vectors, respectively. Obviously

$$Q_{\bar\kappa}(z) = \mathbb{C}^{a,b}(\bar u^C;\bar v^C)\sum_{j=1}^{3}(\kappa_j-1)T_{jj}(z)\mathbb{B}^{a,b}(\bar u^B;\bar v^B). \tag{4.5}$$

Let i be a fixed number from the set $\{1,2,3\}$, and let $\bar\kappa_i = 1$ (that is $\kappa_j = 1$ for $j \ne i$). Then we obtain

$$\mathcal{F}^{(i,i)}_{a,b;\kappa_i}(z|\bar u^C,\bar v^C;\bar u^B,\bar v^B) = \frac{Q_{\bar\kappa}(z)}{\kappa_i - 1}, \tag{4.6}$$

where $\mathcal{F}^{(i,i)}_{a,b;\kappa_i}(z)$ denotes a twisted form factor corresponding to the twist parameters $\bar\kappa_i = 1$.

On the other hand

$$Q_{\bar\kappa}(z) = \big(\tau_{\bar\kappa}(z|\bar u^C;\bar v^C) - \tau(z|\bar u^B;\bar v^B)\big)\ \mathbb{C}^{a,b}(\bar u^C;\bar v^C)\mathbb{B}^{a,b}(\bar u^B;\bar v^B), \tag{4.7}$$

where $\tau_{\bar\kappa}(z|\bar u^C;\bar v^C)$ is the eigenvalue of the twisted transfer matrix on the vector $\mathbb{C}^{a,b}(\bar u^C;\bar v^C)$:

$$\tau_{\bar\kappa}(z|\bar u^C,\bar v^C) = \kappa_1 r_1(z)f(\bar u^C,z) + \kappa_2 f(z,\bar u^C)f(\bar v^C,z) + \kappa_3 r_3(z)f(z,\bar v^C). \tag{4.8}$$

Thus, we arrive at

$$\mathcal{F}^{(i,i)}_{a,b;\kappa_i}(z|\bar u^C,\bar v^C;\bar u^B,\bar v^B) = \frac{\tau_{\bar\kappa}(z|\bar u^C;\bar v^C) - \tau(z|\bar u^B;\bar v^B)}{\kappa_i - 1}\,\mathbb{C}^{a,b}(\bar u^C;\bar v^C)\mathbb{B}^{a,b}(\bar u^B;\bar v^B). \tag{4.9}$$

In the limit $\kappa_i \to 1$ we obtain the usual form factor of $T_{ii}(z)$ [9, 10, 17]

$$\mathcal{F}^{(i,i)}_{a,b}(z|\bar u^C,\bar v^C;\bar u^B,\bar v^B) = \frac{d}{d\kappa_i}\left(\tau_{\bar\kappa}(z|\bar u^C;\bar v^C) - \tau(z|\bar u^B;\bar v^B)\right)\mathbb{C}^{a,b}(\bar u^C;\bar v^C)\mathbb{B}^{a,b}(\bar u^B;\bar v^B)\Big|_{\bar\kappa=1}. \tag{4.10}$$

Here the symbol $\bar\kappa = 1$ means that $\kappa_j = 1$, $j = 1,2,3$. Thus, evaluating twisted and usual form factors is reduced to the computation of the scalar product between the twisted and usual on-shell Bethe vectors.

5 Main results

We obtain determinant representations for the scalar product of the twisted and usual on-shell Bethe vectors in two cases: $\kappa_3/\kappa_1 = 1$ and $\kappa_3/\kappa_1 = q^2$. In both cases κ_2 remains an arbitrary complex number. We denote the corresponding scalar products by $S^{(1)}_{a,b}$ and $S^{(q^2)}_{a,b}$, respectively. In order to describe determinant representations we first introduce for an arbitrary set of variables $\bar w = \{w_1, \dots, w_n\}$

$$\Delta_n(\bar w) = \prod_{j>k}^{n} g(w_j, w_k), \qquad \Delta'_n(\bar w) = \prod_{j<k}^{n} g(w_j, w_k). \tag{5.1}$$

Let also

$$C_h = h(\bar v^C, \bar v^C) h(\bar v^C, \bar u^B) h(\bar u^B, \bar u^B). \tag{5.2}$$

Then we introduce an $a \times (a+b)$ matrix $\mathsf{N}^{(u)}(u^C_j, x_k)$

$$\mathsf{N}^{(u)}(u^C_j, x_k) = (-1)^{a-1} t(u^C_j, x_k) \frac{r_1(x_k)}{f(\bar v^C, x_k)} \frac{h(\bar u^C, x_k)}{h(x_k, \bar u^B)} + \frac{\kappa_2}{\kappa_1} t(x_k, u^C_j) \frac{h(x_k, \bar u^C)}{h(x_k, \bar u^B)}, \tag{5.3}$$

and a $b \times (a+b)$ matrix $\mathsf{N}^{(v)}(v^B_j, x_k)$

$$\mathsf{N}^{(v)}(v^B_j, x_k) = (-1)^{b-1} t(x_k, v^B_j) \frac{r_3(x_k)}{f(x_k, \bar u^B)} \frac{h(x_k, \bar v^B)}{h(\bar v^C, x_k)} + t(v^B_j, x_k) \frac{h(\bar v^B, x_k)}{h(\bar v^C, x_k)}. \tag{5.4}$$

In these formulas

$$\bar x = \{\bar u^B, \bar v^C\} = \{u^B_1, \dots u^B_a, v^C_1, \dots, v^C_b\}. \tag{5.5}$$

Proposition 5.1. *The determinant representation for $S^{(1)}_{a,b}$ has the following form:*

$$S^{(1)}_{a,b} = \mathfrak{p}(\bar v^B) \mathfrak{p}(\bar u^B) C_h \Delta'_a(\bar u^C) \Delta'_b(\bar v^B) \Delta_{a+b}(\bar x) \det_{a+b} \mathsf{N}, \tag{5.6}$$

where

$$\begin{aligned} \mathsf{N}_{j,k} &= \mathsf{N}^{(u)}(u^C_j, x_k), \quad j = 1, \dots, a, \\ \mathsf{N}_{j+a,k} &= \mathsf{N}^{(v)}(v^B_j, x_k), \quad j = 1, \dots, b, \end{aligned} \qquad \{x_1, \dots, x_{a+b}\} = \{u^B_1, \dots u^B_a, v^C_1, \dots, v^C_b\}. \tag{5.7}$$

Comparing this result with the analogous scalar product in the models with $GL(3)$-invariant R-matrix [9] we see that up to the trivial prefactor $\mathfrak{p}(\bar v^B)\mathfrak{p}(\bar u^B)$ it has exactly the same form. One simply should replace the rational functions $g(u,v)$, $f(u,v)$, $h(u,v)$, and $t(u,v)$ by their analogs in $GL(3)$-invariant models:

$$\begin{aligned} g(u,v) &\to g^{(0)}(u,v) = \frac{c}{u-v}, \qquad & f(u,v) &\to f^{(0)}(u,v) = \frac{u-v+c}{u-v}, \\ h(u,v) &\to h^{(0)}(u,v) = \frac{u-v+c}{c}, \qquad & t(u,v) &\to t^{(0)}(u,v) = \frac{c^2}{(u-v)(u-v+c)}, \end{aligned} \tag{5.8}$$

where c is a constant. This correspondence is quite similar to the one that we have for the $GL(2)$ case.

Similarly to (5.7) we define a matrix $\widetilde{\mathsf{N}}$ as

$$
\begin{aligned}
\widetilde{\mathsf{N}}_{j,k} &= x_k \mathsf{N}^{(u)}(u^C_j, x_k), \quad j = 1, \dots, a, \\
\widetilde{\mathsf{N}}_{j+a,k} &= \mathsf{N}^{(v)}(v^B_j, x_k), \quad j = 1, \dots, b.
\end{aligned}
\tag{5.9}
$$

Proposition 5.2. *The determinant representation for $S^{(q^2)}_{a,b}$ has the following form:*

$$
S^{(q^2)}_{a,b} = \mathfrak{p}(\bar v^B) C_h \Delta'_a(\bar u^C) \Delta'_b(\bar v^B) \Delta_{a+b}(\bar x) \det_{a+b} \widetilde{\mathsf{N}}. \tag{5.10}
$$

It is clear that in the scaling limit $q \to 1$ representations (5.10) and (5.6) coincide and give the determinant formula for the scalar product of the twisted and usual on-shell vectors in the models with $GL(3)$-invariant R-matrix [9].

Remark. The matrix elements $\mathsf{N}^{(u)}(u^C_j, x_k)$ (5.3) and $\mathsf{N}^{(v)}(v^B_j, x_k)$ (5.4) require additional definition if $\bar u^C \cap \bar u^B \ne \emptyset$ or $\bar v^C \cap \bar v^B \ne \emptyset$. Let for definiteness $u^C_a = u^B_a$. The matrix element $\mathsf{N}^{(u)}(u^C_a, u^B_a)$ depends on the functions $t(u^C_a, u^B_a)$ and $t(u^B_a, u^C_a)$, which have a pole at $u^B_a = u^C_a$. However one can easily check that the residue in this pole vanishes due to the twisted Bethe equations (4.2), and hence, the limit of (5.3) is finite:

$$
\begin{aligned}
\mathsf{N}^{(u)}(u^C_a, u^B_a)\Big|_{u^C_a = u^B_a} = \frac{\kappa_2}{\kappa_1 u^C_a} \frac{h(u^C_a, \bar u^C)}{h(\bar u^C_a, \bar u^B)} \Bigg[&(q^{-1} - q) \frac{r'_1(u^C_a)}{r_1(u^C_a)} \\
&+ \sum_{\ell=1}^{a-1} \frac{(q + q^{-1}) u^C_\ell}{h(u^C_a, u^C_\ell) h(u^C_\ell, u^C_a)} + \sum_{i=1}^{b} v^C_i t(v^C_i, u^C_a) \Bigg],
\end{aligned}
\tag{5.11}
$$

where $r'_1(u^C_a)$ is the derivative of $r_1(u^C_a)$. It is this expression for $\mathsf{N}^{(u)}(u^C_a, u^B_a)$ that needs to be used in the case $u^C_a = u^B_a$. Similarly, if $v^C_b = v^B_b$, then

$$
\begin{aligned}
\mathsf{N}^{(v)}(v^B_b, v^C_b)\Big|_{v^C_b = v^B_b} = \frac{1}{v^B_b} \frac{h(\bar v^B, v^B_b)}{h(\bar v^C, v^B_b)} \Bigg[&(q - q^{-1}) \frac{r'_3(v^B_b)}{r_3(v^B_b)} \\
&- \sum_{i=1}^{b-1} \frac{(q + q^{-1}) v^B_i}{h(v^B_b, v^B_i) h(v^B_i, v^B_b)} + \sum_{\ell=1}^{a} u^B_\ell t(v^B_b, u^B_\ell) \Bigg],
\end{aligned}
\tag{5.12}
$$

where $r'_3(v^B_b)$ is the derivative of $r_3(v^B_b)$.

Note that in the particular case $\bar u^C = \bar u^B$, $\bar v^C = \bar v^B$, and $\bar\kappa = 1$ representation (5.6) for $S^{(1)}_{a,b}$ gives the square of the norm of the on-shell Bethe vector. In this case one should use (5.11) and (5.12) for the diagonal entries of the matrix N.

We would like to stress once more that in the obtained results the parameter κ_2 is an arbitrary complex number, while the ratio of the parameters κ_1 and κ_3 is fixed. Since the final result depends on the ratios of κ_i only, one can say that κ_1 and κ_3 are fixed. This means that

using (4.10) we can obtain a determinant representation for the form factor of the operator $T_{22}(z)$

$$\mathcal{F}^{(2,2)}_{a,b}(z|\bar u^C,\bar v^C;\bar u^B,\bar v^B)=\frac{d}{d\kappa_2}\left(\tau_{\bar\kappa}(z|\bar u^C;\bar v^C)-\tau(z|\bar u^B;\bar v^B)\right)\,S^{(1)}_{a,b}(\bar u^C,\bar v^C;\bar u^B,\bar v^B)\Big|_{\kappa_2=1}. \tag{5.13}$$

However, we cannot obtain determinant representations for the form factors of the operators $T_{11}(z)$ and $T_{33}(z)$. For these operators we know only the twisted form factors with a fixed values of the twist parameters, for example

$$\mathcal{F}^{(3,3)}_{a,b;q^2}(z|\bar u^C,\bar v^C;\bar u^B,\bar v^B)=\frac{\tau_{\bar\kappa}(z|\bar u^C;\bar v^C)-\tau(z|\bar u^B;\bar v^B)}{q^2-1}\,S^{(q^2)}_{a,b}(\bar u^C,\bar v^C;\bar u^B,\bar v^B), \tag{5.14}$$

where $\bar\kappa=\{1,1,q^2\}$.

The remainder of this work is devoted to the proof of propositions 5.1 and 5.2.

6 Summation identities

The central object of the theory of scalar products in the $GL(2)$-based models is the Izergin determinant $\mathsf{K}_k(\bar x|\bar y)$ [20]. It also plays an important role in the case of the $GL(3)$-based models with trigonometric R-matrix. The Izergin determinant is defined for two sets $\bar x$ and $\bar y$ of the same cardinality $\#\bar x=\#\bar y=k$:

$$\mathsf{K}_k(\bar x|\bar y)=\Delta'(\bar x)\Delta(\bar y)\;h(\bar x,\bar y)\;\det_k t(x_i,y_j)\,. \tag{6.1}$$

For further applications it is convenient to introduce two modifications of the Izergin determinant

$$\mathsf{K}^{(l)}_k(\bar x|\bar y)=\mathfrak{p}(\bar x)\cdot\mathsf{K}_k(\bar x|\bar y)\,,\qquad \mathsf{K}^{(r)}_k(\bar x|\bar y)=\mathfrak{p}(\bar y)\cdot\mathsf{K}_k(\bar x|\bar y)\,, \tag{6.2}$$

which we call left and right Izergin determinants, respectively.

For the derivation of determinant representations for scalar products we shall use several properties of the Izergin determinant. It is obvious that this rational function is symmetric over the set $\bar x$ and symmetric over the set $\bar y$. The Izergin determinant (6.1) vanishes if one of its arguments goes to infinity and other arguments are fixed. Respectively, $\mathsf{K}^{(l)}_k$ is bounded if $x_i\to\infty$, while $\mathsf{K}^{(r)}_k$ is bounded if $y_i\to\infty$.

Several simple properties of the Izergin determinant follow directly from the definition (6.1). Their detailed proofs are given in [7]. A reduction property has the form

$$\mathsf{K}^{(l,r)}_{n+m}(\{\bar x,q^{-2}\bar z\}|\{\bar y,\bar z\})=\mathsf{K}^{(l,r)}_{n+m}(\{\bar x,\bar z\}|\{\bar y,q^2\bar z\})=(-q)^{\mp m}\mathsf{K}^{(l,r)}_{n}(\bar x|\bar y), \tag{6.3}$$

where $m=\#\bar z$ and $n=\#\bar x=\#\bar y$. Here the superscripts (l,r) on K mean that the equation is valid both for $\mathsf{K}^{(l)}$ and for $\mathsf{K}^{(r)}$ with an appropriate choice of component (first/up or second/down) throughout the equation.

A similar reduction of the Izergin determinant takes place in the poles of this function

$$\mathsf{K}^{(l,r)}_{n+1}(\{\bar x,z\}|\{\bar y,z'\})\Big|_{z'\to z}=f(z,z')f(z,\bar y)f(\bar x,z)\mathsf{K}^{(l,r)}_{n}(\bar x|\bar y)+\text{reg}, \tag{6.4}$$

where reg means the regular part. One more simple property of $\mathsf{K}_n^{(l,r)}$ is also useful

$$\mathsf{K}_n^{(l,r)}(q^{-2}\bar{x}|\bar{y}) = \mathsf{K}_n^{(l,r)}(\bar{x}|q^2\bar{y}) = (-q)^{\mp n} f^{-1}(\bar{y},\bar{x})\mathsf{K}_n^{(r,l)}(\bar{y}|\bar{x})\,. \tag{6.5}$$

More sophisticated properties of the Izergin determinant represent summation identities, in which the sum is taken over the partitions of one or more sets of variables.

Lemma 6.1. *Let $\bar{\gamma}$, $\bar{\alpha}$ and $\bar{\beta}$ be three sets of complex variables with $\#\alpha = m_1$, $\#\beta = m_2$, and $\#\gamma = m_1 + m_2$. Then*

$$\sum \mathsf{K}_{m_1}^{(l,r)}(\bar{\gamma}_{\rm I}|\bar{\alpha})\mathsf{K}_{m_2}^{(r,l)}(\bar{\beta}|\bar{\gamma}_{\rm II}) f(\bar{\gamma}_{\rm II},\bar{\gamma}_{\rm I}) = (-q)^{\mp m_1} f(\bar{\gamma},\bar{\alpha})\mathsf{K}_{m_1+m_2}^{(r,l)}(\{\bar{\alpha}q^{-2},\bar{\beta}\}|\bar{\gamma}). \tag{6.6}$$

The sum is taken with respect to all partitions of the set $\bar{\gamma} \Rightarrow \{\bar{\gamma}_{\rm I},\bar{\gamma}_{\rm II}\}$ with $\#\bar{\gamma}_{\rm I} = m_1$ and $\#\bar{\gamma}_{\rm II} = m_2$.

This lemma is a complete analog of lemma 1 of [9], where the reader can find the detailed proof.

Lemma 6.2. *Let $\bar{\gamma}$ and $\bar{\xi}$ be two sets of generic complex numbers with $\#\bar{\gamma} = \#\bar{\xi} = m$. Let also $\phi_1(\gamma)$ and $\phi_2(\gamma)$ be two arbitrary functions of a complex variable γ. Then*

$$\begin{aligned}&\sum \mathsf{K}_m^{(l,r)}(\{\bar{\gamma}_{\rm I}q^{-2},\bar{\gamma}_{\rm II}\}|\bar{\xi}) f(\bar{\xi},\bar{\gamma}_{\rm I}) f(\bar{\gamma}_{\rm II},\bar{\gamma}_{\rm I})\phi_1(\bar{\gamma}_{\rm I})\phi_2(\bar{\gamma}_{\rm II})\\ &\quad= \mathfrak{p}^{\ell,r}\,\Delta'_m(\bar{\xi})\Delta_m(\bar{\gamma})\det_m\Big(\phi_2(\gamma_k)t(\gamma_k,\xi_j)h(\gamma_k,\bar{\xi}) + q^{\mp 1}(-1)^m\phi_1(\gamma_k)t(\xi_j,\gamma_k)h(\bar{\xi},\gamma_k)\Big).\end{aligned} \tag{6.7}$$

Here $\mathfrak{p}^\ell = \mathfrak{p}(\bar{\gamma})$, $\mathfrak{p}^r = \mathfrak{p}(\bar{\xi})$, and we used the shorthand notation (2.9) for the products of the functions ϕ_1 and ϕ_2. The sum is taken over all possible partitions $\bar{\gamma} \Rightarrow \{\bar{\gamma}_{\rm I},\bar{\gamma}_{\rm II}\}$.

This lemma is a trigonometric analog of lemma 2 of [9]. However, the proof contains small, but important differences. Therefore, we give this proof in appendix A.

Lemma 6.3. *Let $\bar{\alpha}$ and $\bar{\beta}$ be two sets of generic complex numbers with $\#\bar{\alpha} = \#\bar{\beta} = n$. Let z be an arbitrary complex number. Then*

$$\sum q^{n_{\rm I}} f(\bar{\beta}_{\rm I},z) f(\bar{\beta}_{\rm II},\bar{\beta}_{\rm I}) f(\bar{\alpha}_{\rm I},\bar{\alpha}_{\rm II})\mathsf{K}_{n_{\rm I}}^{(r)}(\bar{\beta}_{\rm I}|\bar{\alpha}_{\rm I})\mathsf{K}_{n_{\rm II}}^{(l)}(\bar{\alpha}_{\rm II}|\bar{\beta}_{\rm II}q^{-2}) = q^n G_n(\bar{\alpha}|\bar{\beta})h(\bar{\alpha},z)g(\bar{\beta},z), \tag{6.8}$$

where

$$G_n(\bar{\alpha}|\bar{\beta}) = (-1)^n t(\bar{\alpha},\bar{\beta})h(\bar{\alpha},\bar{\alpha})h(\bar{\beta},\bar{\beta}), \tag{6.9}$$

and the sum is taken over all possible partitions $\bar{\alpha} \Rightarrow \{\bar{\alpha}_{\rm I},\ \bar{\alpha}_{\rm II}\}$ and $\bar{\beta} \Rightarrow \{\bar{\beta}_{\rm I},\ \bar{\beta}_{\rm II}\}$ with $\#\bar{\alpha}_{\rm I} = \#\bar{\beta}_{\rm I} = n_{\rm I}$, $n_{\rm I} = 0,\dots,n$, and $\#\bar{\alpha}_{\rm II} = \#\bar{\beta}_{\rm II} = n_{\rm II} = n - n_{\rm I}$.

This lemma is a trigonometric analog of lemma $A.3$ of [11]. Due to the importance of this lemma we give the proof in appendix B.

The identity (6.8) has several particular cases.

Corollary 6.1. *Let $\bar\alpha$ and $\bar\beta$ be two sets of generic complex numbers with $\#\bar\alpha = \#\bar\beta = n$. Then*

$$\sum f(\bar\beta_{\rm II},\bar\beta_{\rm I})f(\bar\alpha_{\rm I},\bar\alpha_{\rm II})\mathsf{K}^{(r)}_{n_{\rm I}}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}^{(l)}_{n_{\rm II}}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2}) = G_n(\bar\alpha|\bar\beta), \tag{6.10}$$

and

$$\sum q^{2n_{\rm II}} f(\bar\beta_{\rm II},\bar\beta_{\rm I})f(\bar\alpha_{\rm I},\bar\alpha_{\rm II})\mathsf{K}^{(l)}_{n_{\rm I}}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}^{(r)}_{n_{\rm II}}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2}) = \frac{\mathfrak{p}(\bar\beta)}{\mathfrak{p}(\bar\alpha)}\, G_n(\bar\alpha|\bar\beta). \tag{6.11}$$

The sum is taken over all possible partitions $\bar\alpha \Rightarrow \{\bar\alpha_{\rm I},\ \bar\alpha_{\rm II}\}$ and $\bar\beta \Rightarrow \{\bar\beta_{\rm I},\ \bar\beta_{\rm II}\}$ with $\#\bar\alpha_{\rm I} = \#\bar\beta_{\rm I} = n_{\rm I}$, $n_{\rm I} = 0,\dots,m$, and $\#\bar\alpha_{\rm II} = \#\bar\beta_{\rm II} = n_{\rm II} = n - n_{\rm I}$.

Identity (6.10) follows from (6.8) in the limit $z \to \infty$. Identity (6.11) follows from (6.10) due to obvious relation

$$\mathsf{K}^{(r)}_{n_{\rm I}}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}^{(l)}_{n_{\rm II}}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2}) = q^{2n_{\rm II}}\frac{\mathfrak{p}(\bar\alpha)}{\mathfrak{p}(\bar\beta)}\,\mathsf{K}^{(l)}_{n_{\rm I}}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}^{(r)}_{n_{\rm II}}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2}). \tag{6.12}$$

Corollary 6.2. *Under the conditions of corollary 6.1*

$$\sum q^{2n_{\rm I}} f(\bar\beta_{\rm II},\bar\beta_{\rm I})f(\bar\alpha_{\rm I},\bar\alpha_{\rm II})\mathsf{K}^{(r)}_{n_{\rm I}}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}^{(l)}_{n_{\rm II}}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2}) = q^{2n}\,\frac{\mathfrak{p}(\bar\alpha)}{\mathfrak{p}(\bar\beta)}\, G_n(\bar\alpha|\bar\beta). \tag{6.13}$$

and

$$\sum f(\bar\beta_{\rm II},\bar\beta_{\rm I})f(\bar\alpha_{\rm I},\bar\alpha_{\rm II})\mathsf{K}^{(l)}_{n_{\rm I}}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}^{(r)}_{n_{\rm II}}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2}) = G_n(\bar\alpha|\bar\beta), \tag{6.14}$$

Identity (6.13) follows from (6.8) in the limit $z = 0$. Identity (6.14) follows from (6.13) and (6.12).

Corollary 6.3. *Under the conditions of corollary 6.1*

$$\sum q^{n_{\rm I}}[n_{\rm I}] f(\bar\beta_{\rm II},\bar\beta_{\rm I})f(\bar\alpha_{\rm I},\bar\alpha_{\rm II})\mathsf{K}^{(r)}_{n_{\rm I}}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}^{(l)}_{n_{\rm II}}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2}) = \frac{q^{2n}\frac{\mathfrak{p}(\bar\alpha)}{\mathfrak{p}(\bar\beta)} - 1}{q - q^{-1}}\, G_n(\bar\alpha|\bar\beta), \tag{6.15}$$

where $[n_{\rm I}]$ means a q-number $[n_{\rm I}] = (q^{n_{\rm I}} - q^{-n_{\rm I}})/(q - q^{-1})$.

This identity is a linear combination of (6.10) and (6.13).

7 Derivation of determinant representation

Derivation of a determinant representation for the scalar product of twisted on-shell and usual on-shell Bethe vectors is quite similar to the one given in [9] for $GL(3)$-invariant models. Nevertheless, in certain cases the q-deformation of the R-matrix leads to an essential difference in the final result. Therefore we give all the details of the derivation.

We start with the sum formula for the scalar product in the $GL(3)$-based models with the trigonometric R-matrix [8]:

$$\begin{aligned} S_{a,b}(\bar u^C;\bar v^C|\bar u^B;\bar v^B) = \sum f(\bar u^B_{\rm II},\bar u^B_{\rm I})f(\bar u^C_{\rm I},\bar u^C_{\rm II})f(\bar v^B_{\rm I},\bar v^B_{\rm II})f(\bar v^C_{\rm II},\bar v^C_{\rm I})f(\bar v^C_{\rm I},\bar u^C_{\rm I})f(\bar v^B_{\rm II},\bar u^B_{\rm II}) \\ \times \frac{r_1(\bar u^C_{\rm II})r_1(\bar u^B_{\rm I})r_3(\bar v^C_{\rm II})r_3(\bar v^B_{\rm I})}{f(\bar v^C,\bar u^C)f(\bar v^B,\bar u^B)}\,\mathsf{Z}^{(l)}_{a-k,n}(\bar u^C_{\rm II};\bar u^B_{\rm II}|\bar v^C_{\rm I};\bar v^B_{\rm I})\,\mathsf{Z}^{(r)}_{k,b-n}(\bar u^B_{\rm I};\bar u^C_{\rm I}|\bar v^B_{\rm II};\bar v^C_{\rm II})\,. \end{aligned} \tag{7.1}$$

The sum in (7.1) is taken over the partitions of the sets $\bar u^C$, $\bar u^B$, $\bar v^C$, and $\bar v^B$

$$\begin{aligned} &\bar u^C \Rightarrow \{\bar u^C_{\rm I},\ \bar u^C_{\rm II}\}, \qquad &&\bar v^C \Rightarrow \{\bar v^C_{\rm I},\ \bar v^C_{\rm II}\},\\ &\bar u^B \Rightarrow \{\bar u^B_{\rm I},\ \bar u^B_{\rm II}\}, \qquad &&\bar v^B \Rightarrow \{\bar v^B_{\rm I},\ \bar v^B_{\rm II}\}. \end{aligned} \tag{7.2}$$

The partitions are independent except that $\#\bar u^B_{\rm I} = \#\bar u^C_{\rm I} = k$ with $k = 0,\dots,a$, and $\#\bar v^B_{\rm I} = \#\bar v^C_{\rm I} = n$ with $n = 0,\dots,b$.

The functions $\mathsf{Z}^{(l)}_{a-k,n}$ and $\mathsf{Z}^{(r)}_{k,b-n}$ are the left and the right highest coefficients, respectively. They can be expressed in terms of the Izergin determinants [7]. We need two representations of these functions. The first one reads

$$\mathsf{Z}^{(l,r)}_{a,b}(\bar t;\bar x|\bar s;\bar y) = (-q)^{\mp b}\sum \mathsf{K}^{(r,l)}_b(\bar s|\bar w_{\rm I}q^2)\mathsf{K}^{(l,r)}_a(\bar w_{\rm II}|\bar t)\mathsf{K}^{(l,r)}_b(\bar y|\bar w_{\rm I})f(\bar w_{\rm I},\bar w_{\rm II}), \tag{7.3}$$

where $\bar w = \{\bar x,\bar s\}$. The sum is taken with respect to partitions of the set $\bar w \Rightarrow \{\bar w_{\rm I},\bar w_{\rm II}\}$ with $\#\bar w_{\rm I} = b$ and $\#\bar w_{\rm II} = a$. Similarly to the formulas of section 6 (see e.g. (6.6), (6.2)) the superscript (l,r) on $\mathsf{Z}_{a,b}$ means that equation (7.3) is valid for $\mathsf{Z}^{(l)}_{a,b}$ and for $\mathsf{Z}^{(r)}_{a,b}$ separately. Choosing the first or the second component in the pair (l,r) and the corresponding (up or down resp.) exponent of $(-q)^{\mp b}$ in this equation, we obtain representations either for $\mathsf{Z}^{(l)}_{a,b}(\bar t;\bar x|\bar s;\bar y)$ or for $\mathsf{Z}^{(r)}_{a,b}(\bar t;\bar x|\bar s;\bar y)$.

The second representation for the highest coefficient has the following form:

$$\begin{aligned}\mathsf{Z}^{(l,r)}_{a,b}(\bar t;\bar x|\bar s;\bar y) &= (-q)^{\mp a}f(\bar y,\bar x)f(\bar s,\bar t)\\ &\times\sum \mathsf{K}^{(r,l)}_a(\bar tq^{-2}|\bar\eta_{\rm I}q^2)\mathsf{K}^{(l,r)}_a(\bar xq^{-2}|\bar\eta_{\rm I})\mathsf{K}^{(l,r)}_b(\bar\eta_{\rm II}|\bar s)f(\bar\eta_{\rm I},\bar\eta_{\rm II}),\end{aligned} \tag{7.4}$$

where $\bar\eta = \{\bar y,\bar tq^{-2}\}$. The sum is taken with respect to partitions of the set $\bar\eta \Rightarrow \{\bar\eta_{\rm I},\bar\eta_{\rm II}\}$ with $\#\bar\eta_{\rm I} = a$ and $\#\bar\eta_{\rm II} = b$.

It is convenient to use one type of representations for $\mathsf{Z}^{(l)}_{a-k,n}$ and another type for $\mathsf{Z}^{(r)}_{k,b-n}$. For example, using (7.3) for $\mathsf{Z}^{(l)}_{a-k,n}(\bar u^C_{\rm II};\bar u^B_{\rm II}|\bar v^C_{\rm I};\bar v^B_{\rm I})$ we obtain

$$\mathsf{Z}^{(l)}_{a-k,n}(\bar u^C_{\rm II};\bar u^B_{\rm II}|\bar v^C_{\rm I};\bar v^B_{\rm I}) = (-q)^{-n}\sum \mathsf{K}^{(r)}_n(\bar v^C_{\rm I}|\bar w_{\rm I}q^2)\mathsf{K}^{(l)}_{a-k}(\bar w_{\rm II}|\bar u^C_{\rm II})\mathsf{K}^{(l)}_n(\bar v^B_{\rm I}|\bar w_{\rm I})f(\bar w_{\rm I},\bar w_{\rm II}), \tag{7.5}$$

with $\bar w = \{\bar u^B_{\rm II},\bar v^C_{\rm I}\}$. Thus, we obtain an additional sum over partitions of the sets $\bar u^B_{\rm II}$ and $\bar v^C_{\rm I}$ into sub-subsets, but we do not create additional sub-subsets of $\bar u^C$ and $\bar v^B$. Using now (7.4) for $\mathsf{Z}^{(r)}_{k,b-n}(\bar u^B_{\rm I};\bar u^C_{\rm I}|\bar v^B_{\rm II};\bar v^C_{\rm II})$ we obtain

$$\begin{aligned}\mathsf{Z}^{(r)}_{k,b-n}(\bar u^B_{\rm I};\bar u^C_{\rm I}|\bar v^B_{\rm II};\bar v^C_{\rm II}) &= (-q)^{k}f(\bar v^C_{\rm II},\bar x)f(\bar s,\bar u^B_{\rm I})\\ &\sum \mathsf{K}^{(l)}_k(\bar u^B_{\rm I}q^{-2}|\bar\eta_{\rm I}q^2)\mathsf{K}^{(r)}_k(\bar u^C_{\rm I}|\bar\eta_{\rm I}q^2)\mathsf{K}^{(r)}_{b-n}(\bar\eta_{\rm II}|\bar v^B_{\rm II})f(\bar\eta_{\rm I},\bar\eta_{\rm II}),\end{aligned} \tag{7.6}$$

with $\bar\eta = \{\bar v^C_{\rm II},\bar u^B_{\rm I}q^{-2}\}$. Thus, we again create new sub-subsets of the sets $\bar u^B$ and $\bar v^C$, but we do not touch the original partitions of the sets $\bar u^C$ and $\bar v^B$. As a result we obtain a possibility to take the sum over partitions $\bar u^C \Rightarrow \{\bar u^C_{\rm I},\ \bar u^C_{\rm II}\}$ and $\bar v^C \Rightarrow \{\bar v^C_{\rm I},\ \bar v^C_{\rm II}\}$. Indeed substituting into

(7.1): (1) representations (7.5) and (7.6); (2) twisted Bethe equations (4.2) for the product of functions $r_1(\bar u^C_{\rm II})$; (3) standard Bethe equations (3.1) for the product of functions $r_3(\bar v^B_{\rm I})$; we obtain

$$S_{a,b}(\bar u^C;\bar v^C|\bar u^B;\bar v^B)=\sum(-q)^{k-n}\left(\frac{\kappa_2}{\kappa_1}\right)^{a-k} r_1(\bar u^B_{\rm I})r_3(\bar v^C_{\rm II})\,f(\bar u^B_{\rm II},\bar u^B_{\rm I})f(\bar v^C_{\rm II},\bar v^C_{\rm I})$$
$$\times f(\bar u^C_{\rm II},\bar u^C_{\rm I})f(\bar v^B_{\rm II},\bar v^B_{\rm I})\mathsf{K}^{(r)}_n(\bar v^C_{\rm I}|\bar w_{\rm I}q^2)\mathsf{K}^{(l)}_{a-k}(\bar w_{\rm II}|\bar u^C_{\rm II})\mathsf{K}^{(l)}_n(\bar v^B_{\rm I}|\bar w_{\rm I})f(\bar w_{\rm I},\bar w_{\rm II})$$
$$\times \mathsf{K}^{(l)}_k(\bar u^B_{\rm I}q^{-2}|\bar\eta_{\rm I}q^2)\mathsf{K}^{(r)}_k(\bar u^C_{\rm I}|\bar\eta_{\rm I}q^2)\mathsf{K}^{(r)}_{b-n}(\bar\eta_{\rm II}|\bar v^B_{\rm II})f(\bar\eta_{\rm I},\bar\eta_{\rm II})\,. \tag{7.7}$$

Here the sum is taken over the partitions (7.2) and additional partitions

$$\begin{aligned}\{\bar u^B_{\rm II},\bar v^C_{\rm I}\}&=\bar w\Rightarrow\{\bar w_{\rm I},\bar w_{\rm II}\},\\ \{\bar v^C_{\rm II},\bar u^B_{\rm I}q^{-2}\}&=\bar\eta\Rightarrow\{\bar\eta_{\rm I},\bar\eta_{\rm II}\}.\end{aligned} \tag{7.8}$$

Now we can apply lemma 6.1 for the summation over the partitions $\bar u^C\Rightarrow\{\bar u^C_{\rm I},\ \bar u^C_{\rm II}\}$ and $\bar v^C\Rightarrow\{\bar v^C_{\rm I},\ \bar v^C_{\rm II}\}$:

$$\sum\mathsf{K}^{(r)}_k(\bar u^C_{\rm I}|\bar\eta_{\rm I}q^2)\mathsf{K}^{(l)}_{a-k}(\bar w_{\rm II}|\bar u^C_{\rm II})f(\bar u^C_{\rm II},\bar u^C_{\rm I})=(-q)^k f^{-1}(\bar\eta_{\rm I},\bar u^C)\,\mathsf{K}^{(l)}_a(\{\bar\eta_{\rm I},\bar w_{\rm II}\}|\bar u^C), \tag{7.9}$$

$$\sum\mathsf{K}^{(l)}_n(\bar v^B_{\rm I}|\bar w_{\rm I})\mathsf{K}^{(r)}_{b-n}(\bar\eta_{\rm II}|\bar v^B_{\rm II})f(\bar v^B_{\rm II},\bar v^B_{\rm I})=(-q)^{-n}\,f(\bar v^B,\bar w_{\rm I})\mathsf{K}^{(r)}_b(\{\bar w_{\rm I}q^{-2},\bar\eta_{\rm II}\}|\bar v^B), \tag{7.10}$$

where we used (2.6). Then we arrive at

$$S_{a,b}(\bar u^C;\bar v^C|\bar u^B;\bar v^B)=\sum(-q)^{2(k-n)}\left(\frac{\kappa_2}{\kappa_1}\right)^{a-k} r_1(\bar u^B_{\rm I})r_3(\bar v^C_{\rm II})$$
$$\times f(\bar u^B_{\rm II},\bar u^B_{\rm I})f(\bar v^C_{\rm II},\bar v^C_{\rm I})\frac{f(\bar v^B,\bar w_{\rm I})}{f(\bar\eta_{\rm I},\bar u^C)}f(\bar w_{\rm I},\bar w_{\rm II})f(\bar\eta_{\rm I},\bar\eta_{\rm II})$$
$$\times\mathsf{K}^{(r)}_n(\bar v^C_{\rm I}|\bar w_{\rm I}q^2)\mathsf{K}^{(l)}_k(\bar u^B_{\rm I}q^{-2}|\bar\eta_{\rm I}q^2)\mathsf{K}^{(l)}_a(\{\bar\eta_{\rm I},\bar w_{\rm II}\}|\bar u^C)\mathsf{K}^{(r)}_b(\{\bar w_{\rm I}q^{-2},\bar\eta_{\rm II}\}|\bar v^B)\,. \tag{7.11}$$

7.1 Summation of sub-partitions

In order to make the next step in the summation over the partitions we should specify the subsets of $\bar w$ and $\bar\eta$. For this we introduce new sub-subsets as follows:

$$\begin{aligned}&\bar\eta_{\rm I}\Rightarrow\{\bar u^B_{\rm ii}q^{-2},\bar v^C_{\rm i}\}, &&\bar u^B_{\rm I}\Rightarrow\{\bar u^B_{\rm i},\bar u^B_{\rm ii}\},\\ &\bar\eta_{\rm II}\Rightarrow\{\bar u^B_{\rm i}q^{-2},\bar v^C_{\rm ii}\}, &&\bar u^B_{\rm II}\Rightarrow\{\bar u^B_{\rm iii},\bar u^B_{\rm iv}\},\\ &\bar w_{\rm I}\Rightarrow\{\bar u^B_{\rm iv},\bar v^C_{\rm iii}\}, &&\bar v^C_{\rm I}\Rightarrow\{\bar v^C_{\rm iii},\bar v^C_{\rm iv}\},\\ &\bar w_{\rm II}\Rightarrow\{\bar u^B_{\rm iii},\bar v^C_{\rm iv}\}, &&\bar v^C_{\rm II}\Rightarrow\{\bar v^C_{\rm i},\bar v^C_{\rm ii}\}.\end{aligned} \tag{7.12}$$

The cardinalities of the sub-subsets above are $\#\bar{u}_j^B = k_j$ and $\#\bar{v}_j^C = n_j$. Evidently $\sum_{j=\mathrm{i}}^{\mathrm{iv}} k_j = a$, $\sum_{j=\mathrm{i}}^{\mathrm{iv}} n_j = b$ and one can easily see that

$$\begin{aligned} &n_{\mathrm{i}} + n_{\mathrm{ii}} = b - n, \qquad && k_{\mathrm{i}} + k_{\mathrm{ii}} = k,\\ &n_{\mathrm{iii}} + n_{\mathrm{iv}} = n, && k_{\mathrm{iii}} + k_{\mathrm{iv}} = a - k,\\ &n_{\mathrm{i}} = k_{\mathrm{i}}, && n_{\mathrm{iv}} = k_{\mathrm{iv}}. \end{aligned} \tag{7.13}$$

Remark. Note that some relations between the cardinalities of the sub-subsets are implicitly shown by the subscripts of the Izergin determinants. For instance, equation (7.11) contains the Izergin determinant $\mathsf{K}_n^{(r)}(\bar{v}_{\mathrm{I}}^C|\bar{w}_{\mathrm{I}}q^2)$. This means that $\#\bar{v}_{\mathrm{I}}^C = \#\bar{w}_{\mathrm{I}} = n$. Using the fact that $\bar{v}_{\mathrm{I}}^C = \{\bar{v}_{\mathrm{iii}}^C, \bar{v}_{\mathrm{iv}}^C\}$ and $\bar{w}_{\mathrm{I}} = \{\bar{u}_{\mathrm{iv}}^B, \bar{v}_{\mathrm{iii}}^C\}$ we conclude that $n_{\mathrm{iii}} + n_{\mathrm{iv}} = n_{\mathrm{iii}} + k_{\mathrm{iv}} = n$. From this we find $n_{\mathrm{iv}} = k_{\mathrm{iv}}$.

Using the new sub-subsets we recast (7.11) as follows:

$$\begin{aligned} S_{a,b} = \sum \left(\frac{\kappa_2}{\kappa_1}\right)^{a-k} &(-q)^{2(k_{\mathrm{i}}+k_{\mathrm{ii}}-n_{\mathrm{iii}}-n_{\mathrm{iv}})} r_1(\bar{u}_{\mathrm{i}}^B) r_1(\bar{u}_{\mathrm{ii}}^B) r_3(\bar{v}_{\mathrm{i}}^C) r_3(\bar{v}_{\mathrm{ii}}^C)\\ &\times f(\bar{u}_{\mathrm{iii}}^B, \bar{u}_{\mathrm{i}}^B) f(\bar{u}_{\mathrm{iii}}^B, \bar{u}_{\mathrm{ii}}^B) f(\bar{u}_{\mathrm{iv}}^B, \bar{u}_{\mathrm{i}}^B) f(\bar{u}_{\mathrm{iv}}^B, \bar{u}_{\mathrm{ii}}^B) f(\bar{u}_{\mathrm{iv}}^B, \bar{u}_{\mathrm{iii}}^B) f(\bar{u}_{\mathrm{ii}}^B, \bar{u}_{\mathrm{i}}^B)\ f(\bar{v}_{\mathrm{i}}^C, \bar{v}_{\mathrm{iii}}^C) f(\bar{v}_{\mathrm{i}}^C, \bar{v}_{\mathrm{iv}}^C)\\ &\times f(\bar{v}_{\mathrm{ii}}^C, \bar{v}_{\mathrm{iii}}^C) f(\bar{v}_{\mathrm{ii}}^C, \bar{v}_{\mathrm{iv}}^C) f(\bar{v}_{\mathrm{i}}^C, \bar{v}_{\mathrm{ii}}^C) f(\bar{v}_{\mathrm{iii}}^C, \bar{v}_{\mathrm{iv}}^C)\ f(\bar{u}_{\mathrm{iv}}^B, \bar{v}_{\mathrm{iv}}^C) f(\bar{v}_{\mathrm{iii}}^C, \bar{u}_{\mathrm{iii}}^B) \frac{f(\bar{v}_{\mathrm{i}}^C, \bar{u}_{\mathrm{i}}^B q^{-2})}{f(\bar{v}_{\mathrm{ii}}^C, \bar{u}_{\mathrm{ii}}^B)}\\ \times \frac{f(\bar{u}^C, \bar{u}_{\mathrm{ii}}^B) f(\bar{v}^B, \bar{u}_{\mathrm{iv}}^B) f(\bar{v}^B, \bar{v}_{\mathrm{iii}}^C)}{f(\bar{v}_{\mathrm{i}}^C, \bar{u}^C)}\ &\mathsf{K}_n^{(r)}(\{\bar{v}_{\mathrm{iii}}^C, \bar{v}_{\mathrm{iv}}^C\}|\{\bar{u}_{\mathrm{iv}}^B q^2, \bar{v}_{\mathrm{iii}}^C q^2\})\ \mathsf{K}_k^{(l)}(\{\bar{u}_{\mathrm{i}}^B q^{-2}, \bar{u}_{\mathrm{ii}}^B q^{-2}\}|\{\bar{u}_{\mathrm{ii}}^B, \bar{v}_{\mathrm{i}}^C q^2\})\\ &\times \mathsf{K}_a^{(l)}(\{\bar{u}_{\mathrm{ii}}^B q^{-2}, \bar{u}_{\mathrm{iii}}^B, \bar{v}_{\mathrm{i}}^C, \bar{v}_{\mathrm{iv}}^C\}|\bar{u}^C)\ \mathsf{K}_b^{(r)}(\{\bar{u}_{\mathrm{i}}^B q^{-2}, \bar{u}_{\mathrm{iv}}^B q^{-2}, \bar{v}_{\mathrm{iii}}^C q^{-2}, \bar{v}_{\mathrm{ii}}^C\}|\bar{v}^B)\,. \end{aligned} \tag{7.14}$$

Here we used $f(xq^{-2}, y) = f(x, yq^2) = f^{-1}(y, x)$.

The next several steps are relatively simple transformations of (7.14). First of all we combine sub-subsets (7.12) into new groups

$$\begin{aligned} &\{\bar{u}_{\mathrm{i}}^B, \bar{u}_{\mathrm{iv}}^B\} = \bar{u}_{\mathrm{I}}^B, \qquad && \{\bar{v}_{\mathrm{i}}^C, \bar{v}_{\mathrm{iv}}^C\} = \bar{v}_{\mathrm{I}}^C,\\ &\{\bar{u}_{\mathrm{ii}}^B, \bar{u}_{\mathrm{iii}}^B\} = \bar{u}_{\mathrm{II}}^B, && \{\bar{v}_{\mathrm{ii}}^C, \bar{v}_{\mathrm{iii}}^C\} = \bar{v}_{\mathrm{II}}^C, \end{aligned} \tag{7.15}$$

and denote $n_{\mathrm{I}} = \#\bar{u}_{\mathrm{I}}^B = \#\bar{v}_{\mathrm{I}}^C$. Pay attention that the subsets $\bar{u}_{\mathrm{I}}^B$, $\bar{u}_{\mathrm{II}}^B$, $\bar{v}_{\mathrm{I}}^C$, and $\bar{v}_{\mathrm{II}}^C$ are not the same as in (7.11). We use, however, the same notation, as we are dealing with the sum over partitions, and therefore it does not matter how we denote separate terms of this sum.

The sum in (7.14) takes the form

$$
\begin{aligned}
S_{a,b} = \sum \left(\frac{\kappa_2}{\kappa_1}\right)^{a-k} & (-q)^{2(k_{\rm i}+k_{\rm ii}-n_{\rm iii}-n_{\rm iv})} r_1(\bar u^B_{\rm i}) r_1(\bar u^B_{\rm ii}) r_3(\bar v^C_{\rm i}) r_3(\bar v^C_{\rm ii}) \\
\times f(\bar u^B_{\rm iii}, \bar u^B_{\rm ii}) f(\bar u^B_{\rm iv}, \bar u^B_{\rm i}) & f(\bar u^B_{\rm iv}, \bar u^B_{\rm II}) f(\bar u^B_{\rm II}, \bar u^B_{\rm i})\; f(\bar v^C_{\rm i}, \bar v^C_{\rm II}) f(\bar v^C_{\rm II}, \bar v^C_{\rm iv}) f(\bar v^C_{\rm i}, \bar v^C_{\rm iv}) f(\bar v^C_{\rm ii}, \bar v^C_{\rm iii}) \\
& \times f(\bar u^B_{\rm iv}, \bar v^C_{\rm iv}) f(\bar v^C_{\rm iii}, \bar u^B_{\rm iii}) \frac{f(\bar v^C_{\rm i}, \bar u^B_{\rm i} q^{-2}) f(\bar u^C, \bar u^B_{\rm ii}) f(\bar v^B, \bar u^B_{\rm iv}) f(\bar v^B, \bar v^C_{\rm iii})}{f(\bar v^C_{\rm ii}, \bar u^B_{\rm ii}) f(\bar v^C_{\rm i}, \bar u^C)} \\
& \times \mathsf{K}^{(r)}_n(\{\bar v^C_{\rm iii}, \bar v^C_{\rm iv}\}|\{\bar u^B_{\rm iv} q^2, \bar v^C_{\rm iii} q^2\})\; \mathsf{K}^{(l)}_k(\{\bar u^B_{\rm i} q^{-2}, \bar u^B_{\rm ii} q^{-2}\}|\{\bar u^B_{\rm ii}, \bar v^C_{\rm i} q^2\}) \\
& \qquad \times \mathsf{K}^{(l)}_a(\{\bar u^B_{\rm ii} q^{-2}, \bar u^B_{\rm iii}, \bar v^C_{\rm I}\}|\bar u^C)\; \mathsf{K}^{(r)}_b(\{\bar u^B_{\rm I} q^{-2}, \bar v^C_{\rm iii} q^{-2}, \bar v^C_{\rm ii}\}|\bar v^B)\,.
\end{aligned}
\tag{7.16}
$$

Simplifying the Izergin determinants $\mathsf{K}^{(r)}_n$ and $\mathsf{K}^{(l)}_k$ via (6.3) and (6.5) we have

$$
\mathsf{K}^{(r)}_n(\{\bar v^C_{\rm iii}, \bar v^C_{\rm iv}\}|\{\bar u^B_{\rm iv} q^2, \bar v^C_{\rm iii} q^2\}) = (-q)^{n_{\rm iii}+n_{\rm iv}} f^{-1}(\bar u^B_{\rm iv}, \bar v^C_{\rm iv}) \mathsf{K}^{(l)}_{n_{\rm iv}}(\bar u^B_{\rm iv}|\bar v^C_{\rm iv})\,,
\tag{7.17}
$$

and

$$
\mathsf{K}^{(l)}_k(\{\bar u^B_{\rm i} q^{-2}, \bar u^B_{\rm ii} q^{-2}\}|\{\bar u^B_{\rm ii}, \bar v^C_{\rm i} q^2\}) = (-q)^{-(k_{\rm i}+k_{\rm ii})} f^{-1}(\bar v^C_{\rm i}, \bar u^B_{\rm i} q^{-2}) \mathsf{K}^{(r)}_{n_{\rm i}}(\bar v^C_{\rm i}|\bar u^B_{\rm i} q^{-2})\,.
\tag{7.18}
$$

After substitution these formulas into (7.16) we arrive at

$$
\begin{aligned}
S_{a,b} = \sum \left(\frac{\kappa_2}{\kappa_1}\right)^{a-k_{\rm i}-k_{\rm ii}} & (-q)^{k_{\rm i}+k_{\rm ii}-n_{\rm iii}-n_{\rm iv}} r_1(\bar u^B_{\rm i}) r_1(\bar u^B_{\rm ii}) r_3(\bar v^C_{\rm i}) r_3(\bar v^C_{\rm ii}) \\
\times f(\bar u^B_{\rm iii}, \bar u^B_{\rm ii}) f(\bar u^B_{\rm iv}, \bar u^B_{\rm i}) & f(\bar u^B_{\rm iv}, \bar u^B_{\rm II}) f(\bar u^B_{\rm II}, \bar u^B_{\rm i})\; f(\bar v^C_{\rm i}, \bar v^C_{\rm II}) f(\bar v^C_{\rm II}, \bar v^C_{\rm iv}) f(\bar v^C_{\rm i}, \bar v^C_{\rm iv}) f(\bar v^C_{\rm ii}, \bar v^C_{\rm iii}) \\
\times & \frac{f(\bar v^C_{\rm iii}, \bar u^B_{\rm iii}) f(\bar u^C, \bar u^B_{\rm ii}) f(\bar v^B, \bar u^B_{\rm iv}) f(\bar v^B, \bar v^C_{\rm iii})}{f(\bar v^C_{\rm ii}, \bar u^B_{\rm ii}) f(\bar v^C_{\rm i}, \bar u^C)}\; \mathsf{K}^{(l)}_{n_{\rm iv}}(\bar u^B_{\rm iv}|\bar v^C_{\rm iv})\; \mathsf{K}^{(r)}_{n_{\rm i}}(\bar v^C_{\rm i}|\bar u^B_{\rm i} q^{-2}) \\
& \qquad \times \mathsf{K}^{(l)}_a(\{\bar u^B_{\rm ii} q^{-2}, \bar u^B_{\rm iii}, \bar v^C_{\rm I}\}|\bar u^C)\; \mathsf{K}^{(r)}_b(\{\bar u^B_{\rm I} q^{-2}, \bar v^C_{\rm iii} q^{-2}, \bar v^C_{\rm ii}\}|\bar v^B)\,.
\end{aligned}
\tag{7.19}
$$

The last preliminary step is to introduce two auxiliary functions

$$
\hat r_1(u^B_k) = \frac{r_1(u^B_k) f(\bar u^B_k, u^B_k)}{f(u^B_k, \bar u^B_k) f(\bar v^B, u^B_k)}, \qquad \hat r_3(v^C_k) = \frac{\kappa_2 r_3(v^C_k) f(v^C_k, \bar v^C_k)}{\kappa_3 f(\bar v^C_k, v^C_k) f(v^C_k, \bar u^C)}.
\tag{7.20}
$$

Observe that if u^B_k satisfies Bethe equations (3.1) and v^C_k satisfies twisted Bethe equations (4.2), then $\hat r_1(u^B_k) = \hat r_3(v^C_k) = 1$. However, we do not use this fact, treating up to some stage the sets $\bar u^B$ and $\bar v^C$ as generic complex numbers. The reason of this will be clear very soon.

Then one can write the products $r_1(\bar u^B_{\rm i})$ and $r_3(\bar v^C_{\rm i})$ as follows:

$$
r_1(\bar u^B_{\rm i}) = \hat r_1(\bar u^B_{\rm i}) \frac{f(\bar u^B_{\rm i}, \bar u^B_{\rm II}) f(\bar u^B_{\rm i}, \bar u^B_{\rm iv})}{f(\bar u^B_{\rm II}, \bar u^B_{\rm i}) f(\bar u^B_{\rm iv}, \bar u^B_{\rm i})} f(\bar v^B, \bar u^B_{\rm i}),
\tag{7.21}
$$

$$r_3(\bar v^C_{\rm i}) = \hat r_3(\bar v^C_{\rm i})\left(\frac{\kappa_2}{\kappa_3}\right)^{n_{\rm i}}\frac{f(\bar v^C_{\rm II},\bar v^C_{\rm i})f(\bar v^C_{\rm iv},\bar v^C_{\rm i})}{f(\bar v^C_{\rm i},\bar v^C_{\rm II})f(\bar v^C_{\rm i},\bar v^C_{\rm iv})}f(\bar v^C_{\rm i},\bar u^C). \tag{7.22}$$

Substituting (7.21) and (7.22) into (7.19) we obtain

$$\begin{aligned} S_{a,b} = \sum \left(\frac{\kappa_2}{\kappa_1}\right)^{a-k_{\rm ii}}\left(\frac{\kappa_1}{\kappa_3}\right)^{n_{\rm i}}(-q)^{k_{\rm i}+k_{\rm ii}-n_{\rm iii}-n_{\rm iv}}f(\bar u^B_{\rm I},\bar u^B_{\rm II})f(\bar v^C_{\rm II},\bar v^C_{\rm I})\\ \times \hat r_1(\bar u^B_{\rm i})\hat r_3(\bar v^C_{\rm i})f(\bar u^B_{\rm i},\bar u^B_{\rm iv})f(\bar v^C_{\rm iv},\bar v^C_{\rm i})\mathsf{K}^{(l)}_{n_{\rm iv}}(\bar u^B_{\rm iv}|\bar v^C_{\rm iv})\ \mathsf{K}^{(r)}_{n_{\rm i}}(\bar v^C_{\rm i}|\bar u^B_{\rm i}q^{-2})\ \frac{f(\bar v^C_{\rm iii},\bar u^B_{\rm iii})}{f(\bar v^C_{\rm ii},\bar u^B_{\rm ii})}\\ \times \mathsf{K}^{(r)}_b(\{\bar u^B_{\rm I}q^{-2},\bar v^C_{\rm iii}q^{-2},\bar v^C_{\rm ii}\}|\bar v^B)r_3(\bar v^C_{\rm ii})f(\bar v^C_{\rm ii},\bar v^C_{\rm iii})f(\bar v^B,\bar u^B_{\rm iv})f(\bar v^B,\bar v^C_{\rm iii})\\ \times \mathsf{K}^{(l)}_a(\{\bar u^B_{\rm ii}q^{-2},\bar u^B_{\rm iii},\bar v^C_{\rm I}|\bar u^C)\ r_1(\bar u^B_{\rm ii})f(\bar u^B_{\rm iii},\bar u^B_{\rm ii})f(\bar u^C,\bar u^B_{\rm ii}).\end{aligned} \tag{7.23}$$

Our goal is to take the sum over partitions $\bar u^B_{\rm II}\Rightarrow\{\bar u^B_{\rm ii},\bar u^B_{\rm iii}\}$ and $\bar v^C_{\rm II}\Rightarrow\{\bar v^C_{\rm ii},\bar v^C_{\rm iii}\}$ via lemma 6.2. For this purpose we present the ratio $f(\bar v^C_{\rm iii},\bar u^B_{\rm iii})/f(\bar v^C_{\rm ii},\bar u^B_{\rm ii})$ in the following form:

$$\frac{f(\bar v^C_{\rm iii},\bar u^B_{\rm iii})}{f(\bar v^C_{\rm ii},\bar u^B_{\rm ii})} = \frac{f(\bar v^C_{\rm II},\bar u^B_{\rm II})}{f(\bar v^C_{\rm II},\bar u^B_{\rm ii})f(\bar v^C_{\rm ii},\bar u^B_{\rm II})} = f(\bar v^C_{\rm II},\bar u^B_{\rm II})\,\frac{f(\bar v^C_{\rm I},\bar u^B_{\rm ii})f(\bar v^C_{\rm ii},\bar u^B_{\rm I})}{f(\bar v^C,\bar u^B_{\rm ii})f(\bar v^C_{\rm ii},\bar u^B)}\,. \tag{7.24}$$

Then (7.23) turns into

$$\begin{aligned} S_{a,b} = \sum \left(\frac{\kappa_1}{\kappa_3}\right)^{n_{\rm i}}(-q)^{2n_{\rm i}}f(\bar v^C_{\rm II},\bar u^B_{\rm II})f(\bar u^B_{\rm I},\bar u^B_{\rm II})f(\bar v^C_{\rm II},\bar v^C_{\rm I})\\ \times \hat r_1(\bar u^B_{\rm i})\hat r_3(\bar v^C_{\rm i})f(\bar u^B_{\rm i},\bar u^B_{\rm iv})f(\bar v^C_{\rm iv},\bar v^C_{\rm i})\mathsf{K}^{(l)}_{n_{\rm iv}}(\bar u^B_{\rm iv}|\bar v^C_{\rm iv})\ \mathsf{K}^{(r)}_{n_{\rm i}}(\bar v^C_{\rm i}|\bar u^B_{\rm i}q^{-2})\\ \times \mathsf{K}^{(r)}_b(\{\bar u^B_{\rm I}q^{-2},\bar v^C_{\rm iii}q^{-2},\bar v^C_{\rm ii}\}|\bar v^B)\frac{(-q)^{n_{\rm ii}-b}r_3(\bar v^C_{\rm ii})}{f(\bar v^C_{\rm ii},\bar u^B)}f(\bar v^C_{\rm ii},\bar v^C_{\rm iii})f(\bar v^B,\bar u^B_{\rm iv})f(\bar v^B,\bar v^C_{\rm iii})f(\bar v^C_{\rm ii},\bar u^B_{\rm I})\\ \times \mathsf{K}^{(l)}_a(\{\bar u^B_{\rm ii}q^{-2},\bar u^B_{\rm iii},\bar v^C_{\rm I}|\bar u^C)\ \frac{(-q)^{k_{\rm ii}}r_1(\bar u^B_{\rm ii})}{f(\bar v^C,\bar u^B_{\rm ii})}\left(\frac{\kappa_2}{\kappa_1}\right)^{a-k_{\rm ii}}f(\bar u^B_{\rm iii},\bar u^B_{\rm ii})f(\bar u^C,\bar u^B_{\rm ii})f(\bar v^C_{\rm I},\bar u^B_{\rm ii}),\end{aligned} \tag{7.25}$$

where we have used $k_{\rm i}=n_{\rm i}$.

The sum over partitions $\bar u^B_{\rm II}\Rightarrow\{\bar u^B_{\rm ii},\bar u^B_{\rm iii}\}$ is only in the last line of (7.25). We can apply lemma 6.2 to this sum if we set

$$\phi_1(\gamma) = -q\frac{r_1(\gamma)}{f(\bar v^C,\gamma)},\qquad \phi_2(\gamma)=\frac{\kappa_2}{\kappa_1}\,. \tag{7.26}$$

Consider the application of lemma 6.2 in more detail. Denote the sum over the partitions $\bar u^B_{\rm II}\Rightarrow\{\bar u^B_{\rm ii},\bar u^B_{\rm iii}\}$ in the last line of (7.25) by U:

$$U = \sum_{\bar u^B_{\rm II}\Rightarrow\{\bar u^B_{\rm ii},\bar u^B_{\rm iii}\}}\mathsf{K}^{(l)}_a(\{\bar u^B_{\rm ii}q^{-2},\bar u^B_{\rm iii},\bar v^C_{\rm I}|\bar u^C)\ \frac{(-q)^{k_{\rm ii}}r_1(\bar u^B_{\rm ii})}{f(\bar v^C,\bar u^B_{\rm ii})}\left(\frac{\kappa_2}{\kappa_1}\right)^{a-k_{\rm ii}}f(\bar u^B_{\rm iii},\bar u^B_{\rm ii})f(\bar u^C,\bar u^B_{\rm ii})f(\bar v^C_{\rm I},\bar u^B_{\rm ii}). \tag{7.27}$$

Let us also write (6.7) with the functions ϕ_1 and ϕ_2 as in (7.26) and $m = a$

$$\mathfrak{p}(\bar\gamma)\mathsf{L}_a(\bar\gamma|\bar\xi) = \sum_{\bar\gamma\Rightarrow\{\bar\gamma_{\rm I},\bar\gamma_{\rm II}\}} \mathsf{K}_a^{(l)}(\{\bar\gamma_{\rm I}q^{-2},\bar\gamma_{\rm II}\}|\bar\xi)(-q)^{m_{\rm I}}\left(\frac{\kappa_2}{\kappa_1}\right)^{m_{\rm II}} \frac{r_1(\bar\gamma_{\rm I})}{f(\bar v^C,\bar\gamma_{\rm I})} f(\bar\xi,\bar\gamma_{\rm I})f(\bar\gamma_{\rm II},\bar\gamma_{\rm I}), \tag{7.28}$$

where $m_{\rm I} = \#\bar\gamma_{\rm I}$, $m_{\rm II} = \#\bar\gamma_{\rm II}$, and

$$\mathsf{L}_a(\bar\gamma|\bar\xi) = \Delta'_a(\bar\xi)\Delta_a(\bar\gamma)\ \det_a\big[\mathsf{N}^{(u)}(\xi_j,\gamma_k)h(\gamma_k,\bar u^B)\big], \tag{7.29}$$

with $\mathsf{N}^{(u)}(\xi_j,\gamma_k)$ defined in (5.3).

Let now $\bar\gamma = \{\bar u^B_{\rm II},\bar v^C_{\rm I}\}$. Observe that $\phi_1(v^C_k) = 0$ due to (7.26). Therefore if $\bar v^C_{\rm I}\cap\gamma_{\rm I}\ne\emptyset$, then the corresponding contribution to (7.28) vanishes. Hence, all the non-vanishing contributions to the sum (7.28) correspond only to such partitions for which $\bar v^C_{\rm I}\subset\gamma_{\rm II}$. Therefore, we can set $\gamma_{\rm I} = \bar u^C_{\rm ii}$ and $\gamma_{\rm II} = \{\bar u^C_{\rm iii},\bar v^C_{\rm I}\}$. Then $m_{\rm I} = k_{\rm ii}$, $m_{\rm II} = a - k_{\rm ii}$, and (7.28) turns into

$$\begin{aligned}\mathfrak{p}(\bar u^B_{\rm II})\mathfrak{p}(\bar v^C_{\rm I})\mathsf{L}_a(\{\bar u^B_{\rm II},\bar v^C_{\rm I}\}|\bar\xi) = \sum_{\bar u^B_{\rm II}\Rightarrow\{\bar u^B_{\rm ii},\bar u^B_{\rm iii}\}} &\mathsf{K}_a^{(l)}(\{\bar u^B_{\rm ii}q^{-2},\bar u^B_{\rm iii},\bar v^C_{\rm I}|\bar\xi)\\ &\times\frac{(-q)^{k_{\rm ii}}r_1(\bar u^B_{\rm ii})}{f(\bar v^C,\bar u^B_{\rm ii})}\left(\frac{\kappa_2}{\kappa_1}\right)^{a-k_{\rm ii}} f(\bar u^B_{\rm iii},\bar u^B_{\rm ii})f(\bar\xi,\bar u^B_{\rm ii})f(\bar v^C_{\rm I},\bar u^B_{\rm ii}).\end{aligned} \tag{7.30}$$

It remains to set $\bar\xi = \bar u^C$, and we reproduce (7.27). Thus, $U = \mathfrak{p}(\bar u^B_{\rm II})\mathfrak{p}(\bar v^C_{\rm I})\mathsf{L}_a(\{\bar u^B_{\rm II},\bar v^C_{\rm I}\}|\bar u^C)$.

Similarly, the sum over partitions $\bar v^C_{\rm II}\Rightarrow\{\bar v^C_{\rm ii},\bar v^C_{\rm iii}\}$ is only in the third line of (7.25). We can apply again lemma 6.2 to this sum with the following identifications:

$$\phi_1(\gamma) = -q^{-1}, \qquad \phi_2(\gamma) = \frac{r_3(\gamma)}{f(\gamma,\bar u^B)}\ . \tag{7.31}$$

Thus, we obtain

$$\begin{aligned}S_{a,b} = \sum f(\bar v^C_{\rm II},\bar u^B_{\rm II})f(\bar u^B_{\rm I},\bar u^B_{\rm II})f(\bar v^C_{\rm II},\bar v^C_{\rm I})\mathfrak{p}(\bar v^B)\mathfrak{p}(\bar u^B_{\rm II})\mathfrak{p}(\bar v^C_{\rm I})&\\ \times\,\mathsf{G}^{(\kappa)}_{n_{\rm I}}(\bar u^B_{\rm I},\bar v^C_{\rm I})\mathsf{L}_a(\{\bar u^B_{\rm II},\bar v^C_{\rm I}\}|\bar u^C)\mathsf{M}_b(\{\bar u^B_{\rm I},\bar v^C_{\rm II}\}|\bar v^B),&\end{aligned} \tag{7.32}$$

where

$$\mathsf{M}_b(\bar\gamma|\bar v^B) = \Delta'_b(\bar v^B)\Delta_b(\bar\gamma)\ \det_b\big[\mathsf{N}^{(v)}(v^B_j,\gamma_k)h(\bar v^C,\gamma_k)\big], \tag{7.33}$$

with $\mathsf{N}^{(v)}(v^B_j,\gamma_k)$ defined in (5.4).

The function $\mathsf{G}^{(\kappa)}_{n_{\rm I}}$ is still given as the sum over the partitions $\bar u^B_{\rm I}\Rightarrow\{\bar u^B_{\rm i},\bar u^B_{\rm iv}\}$ and $\bar v^C_{\rm I}\Rightarrow\{\bar v^C_{\rm i},\bar v^C_{\rm iv}\}$:

$$\mathsf{G}^{(\kappa)}_{n_{\rm I}}(\bar u^B_{\rm I},\bar v^C_{\rm I}) = \sum\left(\frac{\kappa_1}{\kappa_3}\right)^{n_{\rm i}} q^{2n_{\rm i}}\hat r_1(\bar u^B_{\rm i})\hat r_3(\bar v^C_{\rm i})f(\bar u^B_{\rm i},\bar u^B_{\rm iv})f(\bar v^C_{\rm iv},\bar v^C_{\rm i})\mathsf{K}^{(l)}_{n_{\rm iv}}(\bar u^B_{\rm iv}|\bar v^C_{\rm iv})\ \mathsf{K}^{(r)}_{n_{\rm i}}(\bar v^C_{\rm i}|\bar u^B_{\rm i}q^{-2}). \tag{7.34}$$

It is important that up to now we did not use (twisted) Bethe equations for the parameters $\bar v^C$ and $\bar u^B$. In all the above calculations we considered these parameters as generic complex

numbers. Therefore, the matrix elements $\mathsf{N}^{(u)}(u_j^C, u_k^B)$ (5.3) and $\mathsf{N}^{(v)}(v_j^B, v_k^C)$ (5.4) are well defined even in the case when $\bar{u}^C \cap \bar{u}^B \neq \emptyset$ or $\bar{v}^C \cap \bar{v}^B \neq \emptyset$. Indeed, since we have no restriction for $\bar{v}^C$ and $\bar{u}^B$, it is no problem to take the limits $u_k^B \to u_j^C$ in (5.3) or $v_k^C \to v_j^B$ in (5.4). Therefore, if some of the Bethe parameters from different Bethe vectors coincide, then the corresponding matrix elements $\mathsf{N}^{(u)}(u_j^C, u_k^B)$ and $\mathsf{N}^{(v)}(v_j^B, v_k^C)$ should be understood as in (5.11) and (5.12).

8 Final summations

After we have defined the matrices $\mathsf{N}^{(u)}$ and $\mathsf{N}^{(v)}$ for all possible values of the Bethe parameters, we can impose (twisted) Bethe equations for the sets $\bar{v}^C$ and $\bar{u}^B$. Then $\hat{r}_1(u_k^B)\hat{r}_3(v_k^C) = 1$ and the sum in (7.34) turns into

$$\mathsf{G}_{n_{\rm I}}^{(\kappa)}(\bar{u}_{\rm I}^B, \bar{v}_{\rm I}^C) = \sum \left(\frac{\kappa_1}{\kappa_3}\right)^{n_{\rm i}} q^{2n_{\rm i}} f(\bar{u}_{\rm i}^B, \bar{u}_{\rm iv}^B) f(\bar{v}_{\rm iv}^C, \bar{v}_{\rm i}^C) \mathsf{K}_{n_{\rm iv}}^{(l)}(\bar{u}_{\rm iv}^B|\bar{v}_{\rm iv}^C)\, \mathsf{K}_{n_{\rm i}}^{(r)}(\bar{v}_{\rm i}^C|\bar{u}_{\rm i}^B q^{-2}). \tag{8.1}$$

We cannot calculate this sum at generic κ_1/κ_3. However, we know the result for $\mathsf{G}_{n_{\rm I}}^{(\kappa)}$ in two particular cases, namely $\kappa_3/\kappa_1 = 1$ and $\kappa_3/\kappa_1 = q^2$. They are described by corollaries 6.1 and 6.2. Let us denote the corresponding functions $\mathsf{G}_{n_{\rm I}}^{(\kappa)}$ by $\mathsf{G}_{n_{\rm I}}^{(1)}$ and $\mathsf{G}_{n_{\rm I}}^{(q^2)}$, respectively. Then we have

$$\mathsf{G}_{n_{\rm I}}^{(1)} = (-1)^{n_{\rm I}} \frac{\mathfrak{p}(\bar{u}_{\rm I}^B)}{\mathfrak{p}(\bar{v}_{\rm I}^C)}\, t(\bar{v}_{\rm I}^C, \bar{u}_{\rm I}^B) h(\bar{v}_{\rm I}^C, \bar{v}_{\rm I}^C) h(\bar{u}_{\rm I}^B, \bar{u}_{\rm I}^B), \tag{8.2}$$

and

$$\mathsf{G}_{n_{\rm I}}^{(q^2)} = (-1)^{n_{\rm I}}\, t(\bar{v}_{\rm I}^C, \bar{u}_{\rm I}^B) h(\bar{v}_{\rm I}^C, \bar{v}_{\rm I}^C) h(\bar{u}_{\rm I}^B, \bar{u}_{\rm I}^B). \tag{8.3}$$

Consider these two cases separately.

8.1 The case $\kappa_3/\kappa_1 = 1$

Without loss of generality we can set $\kappa_1 = \kappa_3 = 1$. Let us denote the corresponding scalar product by $S_{a,b}^{(1)}$. Substituting (8.2) into (7.32) we obtain

$$\begin{aligned} S_{a,b}^{(1)} = \mathfrak{p}(\bar{v}^B)\mathfrak{p}(\bar{u}^B) \sum (-1)^{n_{\rm I}} f(\bar{v}_{\rm II}^C, \bar{u}_{\rm II}^B) f(\bar{u}_{\rm I}^B, \bar{u}_{\rm II}^B) f(\bar{v}_{\rm II}^C, \bar{v}_{\rm I}^C) t(\bar{v}_{\rm I}^C, \bar{u}_{\rm I}^B) h(\bar{u}_{\rm I}^B, \bar{u}_{\rm I}^B) h(\bar{v}_{\rm I}^C, \bar{v}_{\rm I}^C) \\ \times \mathsf{L}_a(\{\bar{u}_{\rm II}^B, \bar{v}_{\rm I}^C\}|\bar{u}^C) \mathsf{M}_b(\{\bar{u}_{\rm I}^B, \bar{v}_{\rm II}^C\}|\bar{v}^B). \end{aligned} \tag{8.4}$$

Now we introduce

$$\hat{\mathsf{L}}_a(\bar{\gamma}|\bar{u}^C) = \frac{\mathsf{L}_a(\bar{\gamma}|\bar{u}^C)}{h(\bar{\gamma}, \bar{u}^B)}, \qquad \hat{\mathsf{M}}_b(\bar{\gamma}|\bar{v}^B) = \frac{\mathsf{M}_b(\bar{\gamma}|\bar{v}^B)}{h(\bar{v}^C, \bar{\gamma})}, \tag{8.5}$$

and then the equation (8.4) takes the form

$$\begin{aligned} S_{a,b}^{(1)} = \mathfrak{p}(\bar{v}^B)\mathfrak{p}(\bar{u}^B) \sum (-1)^{n_{\rm I}} f(\bar{v}_{\rm II}^C, \bar{u}_{\rm II}^B) f(\bar{u}_{\rm I}^B, \bar{u}_{\rm II}^B) f(\bar{v}_{\rm II}^C, \bar{v}_{\rm I}^C) t(\bar{v}_{\rm I}^C, \bar{u}_{\rm I}^B) h(\bar{u}_{\rm I}^B, \bar{u}_{\rm I}^B) h(\bar{v}_{\rm I}^C, \bar{v}_{\rm I}^C) \\ \times h(\bar{u}_{\rm II}^B, \bar{u}^B) h(\bar{v}_{\rm I}^C, \bar{u}^B) h(\bar{v}^C, \bar{u}_{\rm I}^B) h(\bar{v}^C, \bar{v}_{\rm II}^C) \hat{\mathsf{L}}_a(\{\bar{u}_{\rm II}^B, \bar{v}_{\rm I}^C\}|\bar{u}^C) \hat{\mathsf{M}}_b(\{\bar{u}_{\rm I}^B, \bar{v}_{\rm II}^C\}|\bar{v}^B). \end{aligned} \tag{8.6}$$

It convenient to present all the functions $f(x,y)$ as $f(x,y)=g(x,y)h(x,y)$, all the functions $t(x,y)$ as $t(x,y)=g(x,y)/h(x,y)$, and then to combine separately all the functions $h(x,y)$ and all the functions $g(x,y)$. We obtain

$$S^{(1)}_{a,b}=\mathfrak{p}(\bar v^B)\mathfrak{p}(\bar u^B)C_h\sum g(\bar v^C_{\rm II},\bar u^B_{\rm II})g(\bar u^B_{\rm I},\bar u^B_{\rm II})g(\bar v^C_{\rm II},\bar v^C_{\rm I})g(\bar u^B_{\rm I},\bar v^C_{\rm I})$$
$$\times\,\hat{\mathsf{L}}_a(\{\bar u^B_{\rm II},\bar v^C_{\rm I}\}|\bar u^C)\hat{\mathsf{M}}_b(\{\bar u^B_{\rm I},\bar v^C_{\rm II}\}|\bar v^B),\qquad(8.7)$$

where C_h is given by (5.2). Recall also that

$$\hat{\mathsf{L}}_a(\bar\gamma|\bar u^C)=\Delta'_a(\bar u^C)\Delta_a(\bar\gamma)\,\det_a\mathsf{N}^{(u)}(u^C_j,\gamma_k),\qquad \hat{\mathsf{M}}_b(\bar\gamma|\bar v^B)=\Delta'_b(\bar v^B)\Delta_b(\bar\gamma)\,\det_b\mathsf{N}^{(v)}(v^B_j,\gamma_k),\qquad(8.8)$$

and the matrices $\mathsf{N}^{(u)}(u^C_j,\gamma_k)$, $\mathsf{N}^{(v)}(v^B_j,\gamma_k)$ are given by (5.3), (5.4). It remains to apply the following

Proposition 8.1. *Let N be an $(a+b)\times(a+b)$ matrix and $\bar x=\{x_1,\dots,x_{a+b}\}$ be a set of variables. Let the matrix elements of N have the form*

$$\begin{aligned}\mathsf{N}_{jk}&=\mathsf{N}^{(1)}_j(x_k),\quad j=1,\dots,a\\ \mathsf{N}_{jk}&=\mathsf{N}^{(2)}_j(x_k),\quad j=a+1,\dots,a+b,\end{aligned}\qquad(8.9)$$

where $\mathsf{N}^{(1)}_j(x)$ and $\mathsf{N}^{(2)}_j(x)$ are some functions of x. Then

$$\Delta_{a+b}(\bar x)\det_{a+b}\mathsf{N}=\sum g(\bar x_{\rm II},\bar x_{\rm I})\,\big[\Delta_a(\bar x_{\rm I})\det_a\big(\mathsf{N}^{(1)}_j(x^{\rm I}_k)\big)\big]\cdot\big[\Delta_b(\bar x_{\rm II})\det_b\big(\mathsf{N}^{(2)}_j(x^{\rm II}_k)\big)\big].\qquad(8.10)$$

Here the sum is taken with respect to the partitions $\bar x\Rightarrow\{\bar x_{\rm I},\bar x_{\rm II}\}$ with $\#\bar x_{\rm I}=a$ and $\#\bar x_{\rm II}=b$. The notation $x^{\rm I}_k$ (resp. $x^{\rm II}_k$) means the k-th element of the subset $x_{\rm I}$ (resp. $x_{\rm II}$), and we assume that the elements in every subset are ordered in the natural order.

Proof. Developing $\det_{a+b}\mathsf{N}$ with respect to the first a rows we obtain

$$\det_{a+b}\mathsf{N}=\sum(-1)^{[P_{\rm I,II}]}\det_a\big(\mathsf{N}^{(1)}_j(x^{\rm I}_k)\big)\cdot\det_b\big(\mathsf{N}^{(2)}_j(x^{\rm II}_k)\big),\qquad(8.11)$$

where the sum is taken with respect to the partitions $\bar x\Rightarrow\{\bar x_{\rm I},\bar x_{\rm II}\}$ with $\#\bar x_{\rm I}=a$ and $\#\bar x_{\rm II}=b$. The symbol $[P_{\rm I,II}]$ means the parity of the permutation $P_{\rm I,II}$ mapping the union of the subsets $\{\bar x_{\rm I},\bar x_{\rm II}\}$ into the naturally ordered set $\bar x$: $P_{\rm I,II}(\{\bar x_{\rm I},\bar x_{\rm II}\})=\bar x$. On the other hand the product $\big(\Delta(\bar x)\big)^{-1}$ coincides up to a constant with the Vandermonde determinant of variables $\bar x$. Reordering the rows of this determinant by the permutation $P^{-1}_{\rm I,II}$ we obtain

$$\Delta_{a+b}(\bar x)=(-1)^{[P_{\rm I,II}]}g(\bar x_{\rm II},\bar x_{\rm I})\Delta_a(\bar x_{\rm I})\Delta_b(\bar x_{\rm II}),\qquad(8.12)$$

and hence,

$$(-1)^{[P_{\rm I,II}]}=\frac{g(\bar x_{\rm II},\bar x_{\rm I})\Delta_a(\bar x_{\rm I})\Delta_b(\bar x_{\rm II})}{\Delta_{a+b}(\bar x)}\qquad(8.13)$$

for the arbitrary partition $\bar{x} \Rightarrow \{\bar{x}_{\rm I}, \bar{x}_{\rm II}\}$. Substituting (8.13) into (8.11) we immediately arrive at (8.10). $\square$

It is easy to see that (8.7) is a particular case of (8.10). Indeed, define a set $\bar{x}$ as in (5.5)

$$\bar{x} = \{\bar{u}^B, \bar{v}^C\} = \{u_1^B, \dots u_a^B, v_1^C, \dots, v_b^C\}, \tag{8.14}$$

and let us set in (8.9)

$$\mathsf{N}_j^{(1)}(x) = \mathsf{N}^{(u)}(u_j^C, x), \qquad \mathsf{N}_j^{(2)}(x) = \mathsf{N}^{(v)}(v_j^B, x). \tag{8.15}$$

Then due to proposition 8.1

$$\begin{aligned}
&\Delta'_a(\bar{u}^C)\Delta'_b(\bar{v}^B)\Delta_{a+b}(\bar{x}) \det_{a+b} \mathsf{N} \\
&\quad = \sum g(\bar{x}_{\rm II}, \bar{x}_{\rm I}) \left[\Delta'_a(\bar{u}^C)\Delta_a(\bar{x}_{\rm I}) \det_a\big(\mathsf{N}^{(u)}(u_j^C, x_k^{\rm I})\big)\right] \cdot \left[\Delta'_b(\bar{v}^B)\Delta_b(\bar{x}_{\rm II}) \det_b\big(\mathsf{N}^{(v)}(v_j^B, x_k^{\rm II})\big)\right] \\
&\quad = \sum g(\bar{x}_{\rm II}, \bar{x}_{\rm I}) \hat{\mathsf{L}}_a(\bar{x}_{\rm I}|\bar{u}^C)\hat{\mathsf{M}}_b(\bar{x}_{\rm II}|\bar{v}^B).
\end{aligned} \tag{8.16}$$

If we set $\bar{x}_{\rm I} = \{\bar{u}^B_{\rm II}, \bar{v}^C_{\rm I}\}$ and $\bar{x}_{\rm II} = \{\bar{u}^B_{\rm I}, \bar{v}^C_{\rm II}\}$, then we reproduce the sum in the r.h.s. (8.7). Thus, we finally obtain the result (5.6).

8.2 The case $\kappa_3/\kappa_1 = q^2$

Consider now the case $\kappa_3/\kappa_1 = q^2$. Let us denote the corresponding scalar product by $S_{a,b}^{(q^2)}$. Substituting (8.3) into (7.32) we obtain

$$\begin{aligned}
S_{a,b}^{(q^2)} = \mathfrak{p}(\bar{v}^B) \sum (-1)^{n_{\rm I}} & f(\bar{v}^C_{\rm II}, \bar{u}^B_{\rm II}) f(\bar{u}^B_{\rm I}, \bar{u}^B_{\rm II}) f(\bar{v}^C_{\rm II}, \bar{v}^C_{\rm I}) t(\bar{v}^C_{\rm I}, \bar{u}^B_{\rm I}) h(\bar{u}^B_{\rm I}, \bar{u}^B_{\rm I}) h(\bar{v}^C_{\rm I}, \bar{v}^C_{\rm I}) \\
& \times \mathfrak{p}(\bar{v}^C_{\rm I})\mathfrak{p}(\bar{u}^B_{\rm II}) \mathsf{L}_a(\{\bar{u}^B_{\rm II}, \bar{v}^C_{\rm I}\}|\bar{u}^C) \mathsf{M}_b(\{\bar{u}^B_{\rm I}, \bar{v}^C_{\rm II}\}|\bar{v}^B).
\end{aligned} \tag{8.17}$$

Acting exactly in the same manner as before we obtain

$$\begin{aligned}
S_{a,b}^{(q^2)} = \mathfrak{p}(\bar{v}^B) C_h \sum & g(\bar{v}^C_{\rm II}, \bar{u}^B_{\rm II}) g(\bar{u}^B_{\rm I}, \bar{u}^B_{\rm II}) g(\bar{v}^C_{\rm II}, \bar{v}^C_{\rm I}) g(\bar{u}^B_{\rm I}, \bar{v}^C_{\rm I}) \\
& \times \mathfrak{p}(\bar{v}^C_{\rm I})\mathfrak{p}(\bar{u}^B_{\rm II}) \hat{\mathsf{L}}_a(\{\bar{u}^B_{\rm II}, \bar{v}^C_{\rm I}\}|\bar{u}^C) \hat{\mathsf{M}}_b(\{\bar{u}^B_{\rm I}, \bar{v}^C_{\rm II}\}|\bar{v}^B).
\end{aligned} \tag{8.18}$$

Comparing the obtained equation with (8.7) we see that the difference is very small, namely, we have the additional factor $\mathfrak{p}(\bar{v}^C_{\rm I})\mathfrak{p}(\bar{u}^B_{\rm II})$ in (8.17). This factor can be easily included into determinants. Consider the matrix $\widetilde{\mathsf{N}}$ (5.9). Then due to proposition 8.1 we have

$$\begin{aligned}
\Delta'_a(\bar{u}^C)\Delta'_b(\bar{v}^B)\Delta_{a+b}(\bar{x}) \det_{a+b} \widetilde{\mathsf{N}} &= \sum g(\bar{x}_{\rm II}, \bar{x}_{\rm I}) \left[\Delta'_a(\bar{u}^C)\Delta_a(\bar{x}_{\rm I}) \det_a\big(x_k^{\rm I}\mathsf{N}^{(u)}(u_j^C, x_k^{\rm I})\big)\right] \\
&\qquad \times \left[\Delta'_b(\bar{v}^B)\Delta_b(\bar{x}_{\rm II}) \det_b\big(\mathsf{N}^{(v)}(v_j^B, x_k^{\rm II})\big)\right] \\
= \sum g(\bar{x}_{\rm II}, \bar{x}_{\rm I})\mathfrak{p}(\bar{x}_{\rm I}) & \left[\Delta'_a(\bar{u}^C)\Delta_a(\bar{x}_{\rm I}) \det_a\big(\mathsf{N}^{(u)}(u_j^C, x_k^{\rm I})\big)\right] \left[\Delta'_b(\bar{v}^B)\Delta_b(\bar{x}_{\rm II}) \det_b\big(\mathsf{N}^{(v)}(v_j^B, x_k^{\rm II})\big)\right] \\
&= \sum g(\bar{x}_{\rm II}, \bar{x}_{\rm I})\mathfrak{p}(\bar{x}_{\rm I}) \hat{\mathsf{L}}_a(\bar{x}_{\rm I}|\bar{u}^C)\hat{\mathsf{M}}_b(\bar{x}_{\rm II}|\bar{v}^B).
\end{aligned} \tag{8.19}$$

After specification $\bar{x}_{\rm I} = \{\bar{u}^B_{\rm II}, \bar{v}^C_{\rm I}\}$ and $\bar{x}_{\rm II} = \{\bar{u}^B_{\rm I}, \bar{v}^C_{\rm II}\}$ we reproduce the sum in the r.h.s. of (8.17). Thus, we finally arrive at (5.10).

9 Discussions

We have described the derivation of the determinant formulas for the scalar product of twisted and usual on-shell Bethe vectors in the generalized $GL(3)$-based model with the trigonometric R-matrix. Comparing this derivation with the one given in [9, 10] for the case of the $GL(3)$-invariant R-matrix we see that not only the general line, but even the details are almost the same. The differences are very small and mostly occur in the presence of degrees of the deformation parameter q. However, these negligible differences eventually lead to drastic consequences. Namely, we succeeded in obtaining the determinant representation for the form factor of the operator $T_{22}(z)$ only. The other two form factors of the diagonal entries were unattainable within the framework of our scheme. Indeed, in order to calculate the form factors of $T_{11}(z)$ and $T_{33}(z)$ we should be able to compute κ-derivatives of the function $\mathsf{G}^{(\kappa)}_{n_{\rm I}}$ (8.1) at $\bar\kappa = 1$

$$\frac{d\mathsf{G}^{(\kappa)}_{n_{\rm I}}}{d\kappa_s}\Big|_{\bar\kappa=1} = \pm\sum n_{\rm i}\, q^{2n_{\rm i}} f(\bar u^B_{\rm i}, \bar u^B_{\rm iv}) f(\bar v^C_{\rm iv}, \bar v^C_{\rm i}) \mathsf{K}^{(l)}_{n_{\rm iv}}(\bar u^B_{\rm iv}|\bar v^C_{\rm iv})\, \mathsf{K}^{(r)}_{n_{\rm i}}(\bar v^C_{\rm i}|\bar u^B_{\rm i} q^{-2}), \tag{9.1}$$

where $s = 1, 3$. In the case of the $GL(3)$-invariant R-matrix an analog of this sum over partitions was calculated in [10]. However, in the q-deformed case the sum (9.1) is not described by summation lemma 6.3 or its corollaries. In particular, we have the q-number $[n_{\rm i}]$ in the identity (6.15), while (9.1) contains the usual number $n_{\rm i}$.

Curiously, for the other matrix elements $T_{ij}(z)$ we have a similar situation: we cannot obtain determinant representation for the usual form factors, however it can be done for twisted form factors if the twist parameters are related to the deformation parameter q. In particular, we have found a determinant representation for the twisted form factor of the operator $T_{12}(z)$, provided $\kappa_3/\kappa_1 = q$ and κ_2 is arbitrary. For completeness we describe this result below; however, we do not give the detailed derivation, because its main steps coincide with the ones of [11], while the peculiarities arising due to the q-deformation are already described in the present paper.

Twisted form factor of $T_{12}(z)$ is defined as

$$\mathcal{F}^{(1,2)}_{a,b;\bar\kappa}(z) \equiv \mathcal{F}^{(1,2)}_{a,b;\bar\kappa}(z|\bar u^C, \bar v^C; \bar u^B, \bar v^B) = \mathbb{C}^{a+1,b}(\bar u^C; \bar v^C) T_{12}(z) \mathbb{B}^{a,b}(\bar u^B; \bar v^B), \tag{9.2}$$

where $\mathbb{C}^{a+1,b}(\bar u^C; \bar v^C)$ is a twisted on-shell Bethe vector, while $\mathbb{B}^{a,b}(\bar u^B; \bar v^B)$ is a standard on-shell Bethe vector. Note that in the distinction of the case considered above the number of elements in the set $\bar u^C$ is $a + 1$.

If $\kappa_3/\kappa_1 = q$, then a determinant representation for $\mathcal{F}^{(1,2)}_{a,b;\bar\kappa}(z)$ has the following form:

$$\mathcal{F}^{(1,2)}_{a,b;\bar\kappa}(z) = z\mathfrak{p}(\bar v^B)\mathfrak{p}(\bar u^B) C_h h(\bar v^C, z) h(z, \bar u^B) \Delta'_{a+1}(\bar u^C) \Delta'_b(\bar v^B) \Delta_{a+b+1}(\bar x) \det_{a+b+1} \mathsf{N}^{(12)}, \tag{9.3}$$

where

$$\begin{aligned} &\mathsf{N}^{(12)}_{j,k} = \mathsf{N}^{(u)}(u^C_j, x_k), \qquad j = 1, \dots, a+1, \\ &\mathsf{N}^{(12)}_{j+a+1,k} = \mathsf{N}^{(v)}(v^B_j, x_k), \quad j = 1, \dots, b, \end{aligned} \qquad \{x_1, \dots, x_{a+b+1}\} = \{u^B_1, \dots u^B_a, z, v^C_1, \dots, v^C_b\}, \tag{9.4}$$

The matrix elements $\mathsf{N}^{(v)}(v^B_j, x_k)$ are given by (5.4), and the matrix elements $\mathsf{N}^{(u)}(u^C_j, x_k)$ are given by (5.3), where one should replace $(-1)^{a-1}$ by $(-1)^a$. This replacement is related to the fact that now the set $\bar u^C$ consists of $a + 1$ elements.

It is worth mentioning that if we consider twisted form factors where $\mathbb{C}^{a,b}(\bar{u}^C;\bar{v}^C)$ is a usual on-shell Bethe vector, while $\mathbb{B}^{a,b}(\bar{u}^B;\bar{v}^B)$ is a twisted on-shell Bethe vector, then in order to obtain a determinant representation one should take inverse ratios of the twist parameters. For example, the determinant representations for these twisted form factors of the operators $T_{11}(z)$ and $T_{33}(z)$ exist, if $\kappa_3/\kappa_1 = q^{-2}$. This is not surprising, because one can say that $T_{\bar{\kappa}}(u)$ is the original monodromy matrix, and then $T(u)$ becomes a twisted monodromy matrix with the twist parameters κ_i^{-1}. In other words only the relative twist is important.

This observation makes it impossible to expand over the twisted form factors with the fixed ratios of the twist parameters. Generically twisted form factors can be used for the expansion of complex operators (see. e.g. [17]). For instance, if we need to calculate the matrix element of the product $T_{11}(z_1)T_{33}(z_2)$, then we can expand it into a form factor series as follows:

$$\begin{aligned}&\mathbb{C}^{a,b}(\bar{u}^C;\bar{v}^C)T_{11}(z_1)T_{33}(z_2)\mathbb{B}^{a,b}(\bar{u}^B;\bar{v}^B)\\&\qquad=\sum_{\{\bar{u},\bar{v}\}}\frac{1}{\|\mathbb{B}^{a,b}(\bar{u};\bar{v})\|^2}\,\mathbb{C}^{a,b}(\bar{u}^C;\bar{v}^C)T_{11}(z_1)\mathbb{B}^{a,b}(\bar{u};\bar{v})\cdot\mathbb{C}^{a,b}(\bar{u};\bar{v})T_{33}(z_2)\mathbb{B}^{a,b}(\bar{u}^B;\bar{v}^B).\end{aligned}\tag{9.5}$$

Here the vectors $\mathbb{B}^{a,b}(\bar{u};\bar{v})$ (and their dual $\mathbb{C}^{a,b}(\bar{u};\bar{v})$) are not necessarily usual on-shell Bethe vectors. They might be twisted on-shell Bethe vectors as well. In the last case we are just dealing with the twisted form factors. Suppose that $\mathbb{C}^{a,b}(\bar{u};\bar{v})$ in (9.5) is a twisted on-shell dual Bethe vector with $\kappa_3/\kappa_1 = q^2$. Then we know the determinant representation for this twisted form factor. But in this case the twisted on-shell Bethe vector $\mathbb{B}^{a,b}(\bar{u};\bar{v})$ also has $\kappa_3/\kappa_1 = q^2$, and hence, we do not know a determinant representation for this twisted form factor. Thus, the determinant representations obtained above cannot be used in (9.5).

In conclusion of this discussion we would like to comment on the zero modes method developed in [13] for calculating form factors in $GL(3)$-invariant models. There it was proved that the form factors of all matrix elements $T_{ij}(z)$ can be obtained from one initial form factor, say $\mathcal{F}_{a,b}^{(2,2)}(z)$, by taking the special limits of the Bethe parameters. Since in the q-deformed case the determinant formula for the form factor $\mathcal{F}_{a,b}^{(2,2)}(z)$ is known, one could hope to apply the same approach in order to derive determinants in the q-deformed case for other operators $T_{ij}(z)$. However, the zero modes method fails for the models with the trigonometric R-matrix. It is possible that there exists some modification of this method, which makes it applicable to such models. But to date such a modification is not known.

In the light of the foregoing, the existence of the determinant formula for the form factor of $T_{22}(z)$ appears to be a miraculous exception. At the same time this fact allows us to hope that determinant formulas for other form factors also exist and can be obtained by certain improvement of the method described in the present paper.

Acknowledgements

It is a great pleasure for me to thank my colleagues S. Pakuliak and E. Ragousy for numerous and fruitful discussions. This work was supported by the RSF under a grant 14-50-00005.

A Proof of lemma 6.2

Consider a partition of a set $\bar{\gamma}$ into two subsets $\{\bar{\gamma}_{\rm I}, \bar{\gamma}_{\rm II}\}$ with $\#\bar{\gamma}_{\rm I} = n$, $n = 0, \dots, m$. Let us write explicitly the Izergin determinant $\mathsf{K}_m^{(l,r)}(\{\bar{\gamma}_{\rm I} q^{-2}, \bar{\gamma}_{\rm II}\}|\bar{\xi})$:

$$\mathsf{K}_m^{(l,r)}(\{\bar{\gamma}_{\rm I} q^{-2}, \bar{\gamma}_{\rm II}\}|\bar{\xi}) = q^{-n\mp n}\mathfrak{p}^{\ell,r}\Delta'_m(\bar{\xi})\Delta_n(\bar{\gamma}_{\rm I} q^{-2})\Delta_{m-n}(\bar{\gamma}_{\rm II})\, g(\bar{\gamma}_{\rm II}, \bar{\gamma}_{\rm I} q^{-2}) \\ \times h(\bar{\gamma}_{\rm I} q^{-2}, \bar{\xi}) h(\bar{\gamma}_{\rm II}, \bar{\xi}) \det_m \Big(t(\gamma_k^{\rm I} q^{-2}, \xi_j) \Big| t(\gamma_k^{\rm II}, \xi_j) \Big). \quad \text{(A.1)}$$

Here the symbols $\gamma_k^{\rm I}$ (resp. $\gamma_k^{\rm II}$) denote the k-th element of the subset $\gamma_{\rm I}$ (resp. $\gamma_{\rm II}$). Recall that the elements in every subset are ordered in the natural order. The matrix elements in the first n columns of the determinant in (A.1) are equal to $t(\gamma_k^{\rm I} q^{-2}, \xi_j)$, while in the last $(m-n)$ columns we have $t(\gamma_k^{\rm II}, \xi_j)$. Due to relations (2.6) the equation (A.1) can be written in the form

$$\mathsf{K}_m^{(l,r)}(\{\bar{\gamma}_{\rm I} q^{-2}, \bar{\gamma}_{\rm II}\}|\bar{\xi}) = q^{\mp n}\mathfrak{p}^{\ell,r} \frac{\Delta'_m(\bar{\xi})\Delta_n(\bar{\gamma}_{\rm I})\Delta_{m-n}(\bar{\gamma}_{\rm II}) h(\bar{\gamma}_{\rm II}, \bar{\xi})}{h(\bar{\gamma}_{\rm II}, \bar{\gamma}_{\rm I}) g(\bar{\gamma}_{\rm I}, \bar{\xi})} \det_m \Big(t(\xi_j, \gamma_k^{\rm I}) \Big| t(\gamma_k^{\rm II}, \xi_j) \Big). \quad \text{(A.2)}$$

Hence, we obtain

$$\mathsf{K}_m^{(l,r)}(\{\bar{\gamma}_{\rm I} q^{-2}, \bar{\gamma}_{\rm II}\}|\bar{\xi}) f(\bar{\gamma}_{\rm II}, \bar{\gamma}_{\rm I}) f(\bar{\xi}, \bar{\gamma}_{\rm I}) = \mathfrak{p}^{\ell,r}\Delta'_m(\bar{\xi})\Delta_n(\bar{\gamma}_{\rm I})\Delta_{m-n}(\bar{\gamma}_{\rm II}) g(\bar{\gamma}_{\rm II}, \bar{\gamma}_{\rm I}) \\ \times \det_m \Big((-1)^m q^{\mp 1} t(\xi_j, \gamma_k^{\rm I}) h(\bar{\xi}, \gamma_k^{\rm I}) \;\Big|\; t(\gamma_k^{\rm II}, \xi_j) h(\gamma_k^{\rm II}, \bar{\xi}) \Big). \quad \text{(A.3)}$$

Consider now the r.h.s. of (6.7) as a functional of functions $\phi_1(\gamma)$ and $\phi_2(\gamma)$. Clearly this functional is linear with respect to any $\phi_1(\gamma_k)$ and $\phi_2(\gamma_k)$, $k = 1, \dots, m$. Therefore we have

$$\mathfrak{p}^{\ell,r}\; \Delta'_m(\bar{\xi})\Delta_m(\bar{\gamma}) \det_m \Big(\phi_2(\gamma_k) t(\gamma_k, \xi_j) h(\gamma_k, \bar{\xi}) + q^{\mp 1}(-1)^m \phi_1(\gamma_k) t(\xi_j, \gamma_k) h(\bar{\xi}, \gamma_k) \Big) \\ = \sum \phi_1(\bar{\gamma}_{\rm I}) \phi_2(\bar{\gamma}_{\rm II})\; A(\bar{\gamma}_{\rm I}, \bar{\gamma}_{\rm II}), \quad \text{(A.4)}$$

where the sum is taken over all the partitions $\bar{\gamma} \Rightarrow \{\bar{\gamma}_{\rm I}, \bar{\gamma}_{\rm II}\}$, and the coefficients $A(\bar{\gamma}_{\rm I}, \bar{\gamma}_{\rm II})$ do not depend on ϕ_1 and ϕ_2. In order to find $A(\bar{\gamma}_{\rm I}, \bar{\gamma}_{\rm II})$ corresponding to fixed subsets $\bar{\gamma}_{\rm I}$ and $\bar{\gamma}_{\rm II}$ we should set in the l.h.s. of (A.4)

$$\begin{aligned} \phi_1(\gamma_k) &= 0, \quad \text{if} \quad \gamma_k \in \bar{\gamma}_{\rm II}, \\ \phi_2(\gamma_k) &= 0, \quad \text{if} \quad \gamma_k \in \bar{\gamma}_{\rm I}. \end{aligned} \quad \text{(A.5)}$$

Due to the symmetry of the l.h.s. of (A.4) with respect to all γ_k we can move all columns corresponding to the elements $\gamma_k \in \bar{\gamma}_{\rm I}$ to the left. Simultaneously we should make the same reordering in the product $\Delta_m(\bar{\gamma})$. Then we obtain

$$A(\bar{\gamma}_{\rm I}, \bar{\gamma}_{\rm II}) = \mathfrak{p}^{\ell,r}\Delta'_m(\bar{\xi})\Delta_n(\bar{\gamma}_{\rm I})\Delta_{m-n}(\bar{\gamma}_{\rm II}) g(\bar{\gamma}_{\rm II}, \bar{\gamma}_{\rm I}) \\ \times \det_m \Big((-1)^m q^{\mp 1} t(\xi_j, \gamma_k^{\rm I}) h(\bar{\xi}, \gamma_k^{\rm I}) \;\Big|\; t(\gamma_k^{\rm II}, \xi_j) h(\gamma_k^{\rm II}, \bar{\xi}) \Big). \quad \text{(A.6)}$$

Comparing this equation with (A.3) we arrive at the statement of lemma. □

B Proof of lemma 6.3

The proof is based on induction over n. It is convenient to denote the l.h.s. and the r.h.s. of (6.8) respectively by $\Lambda_n^{(l)}(\bar\alpha|\bar\beta)$ and $\Lambda_n^{(r)}(\bar\alpha|\bar\beta)$. Then we consider the residues of these rational functions in the poles. The last one occurs in the points $\alpha_i=\beta_j$ and $\alpha_i=\beta_j q^{-2}$, $i,j=1,\dots,n$. The goal is to prove that the residues in the above poles can be expressed in terms of $\Lambda_{n-1}^{(l)}(\bar\alpha|\bar\beta)$ and $\Lambda_{n-1}^{(r)}(\bar\alpha|\bar\beta)$. This will give us the induction step.

First of all observe that at $n_{\rm I}$ and $n_{\rm II}$ fixed the sum over partitions of the set $\bar\alpha$ can be calculated explicitly via lemma 6.1

$$\sum_{\bar\alpha\Rightarrow\{\bar\alpha_{\rm I},\,\bar\alpha_{\rm II}\}} f(\bar\alpha_{\rm I},\bar\alpha_{\rm II})\mathsf{K}_{n_{\rm I}}^{(r)}(\bar\beta_{\rm I}|\bar\alpha_{\rm I})\mathsf{K}_{n_{\rm II}}^{(l)}(\bar\alpha_{\rm II}|\bar\beta_{\rm II}q^{-2})=(-q)^{-n_{\rm II}}f(\bar\alpha,\bar\beta_{\rm II}q^{-2})\mathsf{K}_n^{(l)}(\{\bar\beta_{\rm II}q^{-4},\bar\beta_{\rm I}\}|\bar\alpha). \tag{B.1}$$

Thus, the l.h.s. of (6.8) takes the form

$$\Lambda_n^{(l)}(\bar\alpha|\bar\beta)=\sum_{\bar\beta\Rightarrow\{\bar\beta_{\rm I},\,\bar\beta_{\rm II}\}}(-1)^{n_{\rm II}}q^{n_{\rm I}-n_{\rm II}}f(\bar\beta_{\rm I},z)f(\bar\beta_{\rm II},\bar\beta_{\rm I})f(\bar\alpha,\bar\beta_{\rm II}q^{-2})\mathsf{K}_n^{(l)}(\{\bar\beta_{\rm II}q^{-4},\bar\beta_{\rm I}\}|\bar\alpha). \tag{B.2}$$

We will consider $\Lambda_n^{(l)}(\bar\alpha|\bar\beta)$ and $\Lambda_n^{(r)}(\bar\alpha|\bar\beta)$ as functions of α_n and other fixed variables. They are rational functions of α_n and they have poles at[3] $\alpha_n=\beta_k$ and $\alpha_n=\beta_k q^{-2}$, $k=1,\dots,n$. Consider the properties of these functions in the mentioned poles. Due to the symmetry of both functions over $\bar\beta$ it is enough to consider the case $\beta_k=\beta_n$.

Let $\alpha_n=\beta_n$. The pole of $\Lambda_n^{(l)}(\bar\alpha|\bar\beta)$ occurs if and only if $\beta_n\in\bar\beta_{\rm I}$. Let $\bar\beta_{\rm I}=\{\beta_n,\bar\beta_{\rm I'}\}$. Then setting $n_{\rm I'}=n_{\rm I}-1$ and using the property of the Izergin determinant (6.4) we obtain

$$\begin{aligned}\Lambda_n^{(l)}(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n}&=q\sum_{\bar\beta\Rightarrow\{\bar\beta_{\rm I'},\,\bar\beta_{\rm II}\}}(-1)^{n_{\rm II}}q^{n_{\rm I'}-n_{\rm II}}f(\beta_n,z)f(\bar\beta_{\rm I'},z)f(\bar\beta_{\rm II},\beta_n)f(\bar\beta_{\rm II},\bar\beta_{\rm I'})\\ &\quad\times f(\alpha_n,\bar\beta_{\rm II}q^{-2})f(\bar\alpha_n,\bar\beta_{\rm II}q^{-2})f(\beta_n,\alpha_n)f(\alpha_n,\bar\alpha_n)f(\bar\beta_{\rm I'},\beta_n)f(\bar\beta_{\rm II}q^{-4},\beta_n)\\ &\quad\times\mathsf{K}_{n-1}^{(l)}(\{\bar\beta_{\rm II}q^{-4},\bar\beta_{\rm I'}\}|\bar\alpha_n)+\text{reg}.\end{aligned} \tag{B.3}$$

The terms $f(\beta_n,z)$ and $f(\beta_n,\alpha_n)f(\alpha_n,\bar\alpha_n)$ can be moved out off the sum. The terms $f(\bar\beta_{\rm II},\beta_n)$ and $f(\bar\beta_{\rm I'},\beta_n)$ combine into $f(\bar\beta_n,\beta_n)$ and also can be moved out off the sum. Finally, the terms $f(\alpha_n,\bar\beta_{\rm II}q^{-2})$ and $f(\bar\beta_{\rm II}q^{-4},\beta_n)$ cancel each other. We obtain

$$\begin{aligned}\Lambda_n^{(l)}(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n}&=qf(\beta_n,\alpha_n)f(\beta_n,z)f(\bar\beta_n,\beta_n)f(\alpha_n,\bar\alpha_n)\\ \times\sum_{\bar\beta\Rightarrow\{\bar\beta_{\rm I'},\,\bar\beta_{\rm II}\}}&(-1)^{n_{\rm II}}q^{n_{\rm I'}-n_{\rm II}}f(\bar\beta_{\rm I'},z)f(\bar\beta_{\rm II},\bar\beta_{\rm I'})f(\bar\alpha_n,\bar\beta_{\rm II}q^{-2})\mathsf{K}_{n-1}^{(r)}(\{\bar\beta_{\rm II}q^{-4},\bar\beta_{\rm I'}\}|\bar\alpha_n)+\text{reg}.\end{aligned} \tag{B.4}$$

[3]The poles of the Izergin determinant at $\alpha_n=\beta_k q^{-4}$ are compensated by the zeros of the product $f(\bar\alpha,\bar\beta_{\rm II}q^{-2})$.

The remaining sum evidently gives $\Lambda^{(l)}_{n-1}(\bar\alpha_n|\bar\beta_n)$, and we arrive at

$$\Lambda^{(l)}_n(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n} = qf(\beta_n,\alpha_n)f(\beta_n,z)f(\bar\beta_n,\beta_n)f(\alpha_n,\bar\alpha_n)\Lambda^{(l)}_{n-1}(\bar\alpha_n|\bar\beta_n) + \text{reg}. \tag{B.5}$$

Let now $\alpha_n = \beta_n q^{-2}$. The pole of $\Lambda^{(l)}_n(\bar\alpha|\bar\beta)$ occurs if and only if $\beta_n \in \bar\beta_{\rm II}$. Let $\bar\beta_{\rm II} = \{\beta_n, \bar\beta_{\rm II'}\}$. Then setting $n_{\rm II'} = n_{\rm II} - 1$ and using the property of the Izergin determinant (6.3)

$$\mathsf{K}^{(l)}_{n+1}(\{\bar x, q^{-2}z\}|\{\bar y, z\}) = -q\mathsf{K}^{(l)}_n(\bar x|\bar y), \tag{B.6}$$

we obtain

$$\Lambda^{(l)}_n(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n q^{-2}} = \sum_{\bar\beta\Rightarrow\{\bar\beta_{\rm I},\,\bar\beta_{\rm II'}\}} (-1)^{n_{\rm II'}} q^{n_{\rm I}-n_{\rm II'}} f(\bar\beta_{\rm I},z)f(\beta_n,\bar\beta_{\rm I})f(\bar\beta_{\rm II'},\bar\beta_{\rm I})$$

$$\times f(\alpha_n,\beta_n q^{-2})f(\bar\alpha_n,\alpha_n)f(\beta_n,\bar\beta_{\rm II'})f(\bar\alpha_n,\bar\beta_{\rm II'}q^{-2})\mathsf{K}^{(l)}_{n-1}(\{\bar\beta_{\rm II'}q^{-4},\bar\beta_{\rm I}\}|\bar\alpha_n) + \text{reg}. \tag{B.7}$$

The terms $f(\alpha_n,\beta_n q^{-2})f(\bar\alpha_n,\alpha_n)$ can be moved out off the sum. The terms $f(\beta_n,\bar\beta_{\rm I})$ and $f(\beta_n,\bar\beta_{\rm II'})$ combine into $f(\beta_n,\bar\beta_n)$ and also can be moved out off the sum. The remaining sum evidently gives $\Lambda^{(l)}_{n-1}(\bar\alpha_n|\bar\beta_n)$, and we arrive at

$$\Lambda^{(l)}_n(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n q^{-2}} = f(\alpha_n,\beta_n q^{-2})f(\bar\alpha_n,\alpha_n)f(\beta_n,\bar\beta_n)\Lambda^{(l)}_{n-1}(\bar\alpha_n|\bar\beta_n) + \text{reg}. \tag{B.8}$$

Consider now the properties of $\Lambda^{(r)}_n(\bar\alpha|\bar\beta)$ in the same points. Obviously

$$\Lambda^{(r)}_n(\bar\alpha|\bar\beta) = -q\alpha_n\beta_n t(\alpha_n,\beta_n)h(\alpha_n,z)g(\beta_n,z)t(\alpha_n,\bar\beta_n)t(\bar\alpha_n,\beta_n)$$

$$h(\alpha_n,\bar\alpha_n)h(\bar\alpha_n,\alpha_n)h(\beta_n,\bar\beta_n)h(\bar\beta_n,\beta_n)\;\Lambda^{(r)}_{n-1}(\bar\alpha_n|\bar\beta_n), \tag{B.9}$$

where we have used $h(x,x) = x$. Taking the limit $\alpha_n \to \beta_n$ we obtain

$$\Lambda^{(r)}_n(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n} = -q\alpha_n^2\frac{g(\alpha_n,\beta_n)}{h(\alpha_n,\alpha_n)}h(\beta_n,z)g(\beta_n,z)\frac{g(\beta_n,\bar\beta_n)}{h(\beta_n,\bar\beta_n)}\frac{g(\bar\alpha_n,\alpha_n)}{h(\bar\alpha_n,\alpha_n)}$$

$$\times h(\alpha_n,\bar\alpha_n)h(\bar\alpha_n,\alpha_n)h(\beta_n,\bar\beta_n)h(\bar\beta_n,\beta_n)\;\Lambda^{(r)}_{n-1}(\bar\alpha_n|\bar\beta_n) + \text{reg}$$

$$= qf(\beta_n,\alpha_n)f(\beta_n,z)f(\bar\beta_n,\beta_n)f(\alpha_n,\bar\alpha_n)\Lambda^{(r)}_{n-1}(\bar\alpha_n|\bar\beta_n) + \text{reg}. \tag{B.10}$$

Similarly, setting $\alpha_n \to \beta_n q^{-2}$ we obtain

$$\Lambda^{(r)}_n(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n q^{-2}} = -q^{-1}\beta_n^2\frac{g(\beta_n q^{-2},\beta_n)}{h(\alpha_n,\beta_n)}h(\beta_n q^{-2},z)g(\beta_n,z)t(\beta_n q^{-2},\bar\beta_n)t(\bar\alpha_n,\alpha_n q^2)$$

$$\times h(\alpha_n,\bar\alpha_n)h(\bar\alpha_n,\alpha_n)h(\beta_n,\bar\beta_n)h(\bar\beta_n,\beta_n)\;\Lambda^{(r)}_{n-1}(\bar\alpha_n|\bar\beta_n) + \text{reg}. \tag{B.11}$$

Using (2.6) and

$$\frac{q^{-1}\beta_n}{h(\alpha_n,\beta_n)} = f(\alpha_n,\beta_n q^{-2})\Big|_{\alpha_n=\beta_n q^{-2}}, \tag{B.12}$$

we immediately recast (B.11) as follows

$$\Lambda_n^{(r)}(\bar\alpha|\bar\beta)\Big|_{\alpha_n=\beta_n q^{-2}} = f(\alpha_n,\beta_n q^{-2}) f(\bar\alpha_n,\alpha_n) f(\beta_n,\bar\beta_n)\Lambda_{n-1}^{(r)}(\bar\alpha_n|\bar\beta_n) + \text{reg}. \tag{B.13}$$

Now everything is ready for induction over n. Assume that

$$\Lambda_{n-1}^{(l)}(\bar\alpha_n|\bar\beta_n) = \Lambda_{n-1}^{(r)}(\bar\alpha_n|\bar\beta_n). \tag{B.14}$$

It is straightforward to check that (B.14) does hold for $n = 2$. Due to the induction assumption and due to equations (B.5), (B.8), (B.10), and (B.13) the difference $\Lambda_n^{(l)}(\bar\alpha|\bar\beta) - \Lambda_n^{(r)}(\bar\alpha|\bar\beta)$ is a bounded function of α_n in the whole complex plain. It is also easy to see that this function is bounded as $\alpha_n \to \infty$. Hence, it is a constant with respect to α_n. But this constant vanishes at $\alpha_n = 0$, hence, $\Lambda_n^{(l)}(\bar\alpha|\bar\beta) - \Lambda_n^{(r)}(\bar\alpha|\bar\beta) = 0$. □

References

[1] L. D. Faddeev, E. K. Sklyanin and L. A. Takhtajan, *Quantum Inverse Problem. I*, Theor. Math. Phys. **40** (1979) 688–706.

[2] L.D. Faddeev, in: Les Houches Lectures *Quantum Symmetries*, eds A. Connes et al, North Holland, (1998) 149.

[3] L. D. Faddeev and L. A. Takhtajan, *The quantum method of the inverse problem and the Heisenberg XYZ model*, Russian Math. Surveys **34** (1979) 11–68.

[4] P. P. Kulish, E. K. Sklyanin, *Quantum inverse scattering method and the Heisenberg ferromagnet*, Phys. Lett. **70** (1979) 461–463;

[5] P. P. Kulish, E. K. Sklyanin, *Solutions of the Yang-Baxter equation*, Zap. Nauchn. Sem. LOMI **95** (1980) 129–160; J. Sov. Math. **19**:5 (1982) 1596–1620 (Engl. transl.).

[6] J. H. H. Perk and C. L. Schultz, *New families of commuting transfer-matrices in Q-state vertex models*, Phys. Lett. A **84** (1981) 407–410.

[7] S. Pakuliak, E. Ragoucy, N. A. Slavnov, *Scalar products in models with $GL(3)$ trigonometric R-matrix. Highest coefficient*, Theor. Math. Phys., **178**:3 (2014) 314–335, arXiv:1311.3500.

[8] S. Pakuliak, E. Ragoucy, N. A. Slavnov, *Scalar products in models with $GL(3)$ trigonometric R-matrix. General case*, Theor. Math. Phys., **180**:1 (2014) 795–814, arXiv:1401.4355.

[9] S. Belliard, S. Pakuliak, E. Ragoucy, N. A. Slavnov, *The algebraic Bethe ansatz for scalar products in $SU(3)$-invariant integrable models*, J. Stat. Mech. (2012) P10017, arXiv:1207.0956.

[10] S. Belliard, S. Pakuliak, E. Ragoucy, N. A. Slavnov, *Form factors in $SU(3)$-invariant integrable models*, J. Stat. Mech. (2013) P04033, `arXiv:1211.3968`.

[11] S. Pakuliak, E. Ragoucy, N. A. Slavnov, *Form factors in quantum integrable models with GL(3)-invariant R-matrix*, Nucl. Phys. B, **881** (2014) 343–368, `arXiv:1312.1488`.

[12] S. Pakuliak, E. Ragoucy, N. A. Slavnov, *Determinant representations for form factors in quantum integrable models with the $GL(3)$-invariant R-matrix*, Theor. Math. Phys. **181**:3 (2014) 1566–1584, `arXiv:1406.5125`.

[13] S. Pakuliak, E. Ragoucy, N. A. Slavnov, *Zero modes method and form factors in quantum integrable models*, Nucl. Phys. B, **893** (2015) 459–481, `arXiv:1412.6037`.

[14] N. A. Slavnov, *Calculation of scalar products of wave functions and form factors in the framework of the algebraic Bethe ansatz*, Theor. Math. Phys. **79**:2 (1989) 502–508.

[15] N. Yu. Reshetikhin, *Calculation of the norm of Bethe vectors in models with $SU(3)$-symmetry*, Zap. Nauchn. Sem. LOMI **150** (1986) 196–213; J. Sov. Math. **46**:1 (1989) 1694–1706 (Engl. transl.).

[16] S. Belliard, S. Pakuliak, E. Ragoucy, N. A. Slavnov, *Bethe vectors of quantum integrable models with $GL(3)$ trigonometric R-matrix*, SIGMA **9** (2013) 058, `arXiv:1304.7602`.

[17] N. Kitanine, J. M. Maillet, N. A. Slavnov and V. Terras, *Master equation for spin-spin correlation functions of the XXZ chain*, Nucl. Phys. B **712** (2005) 600–622, `arXiv:hep-th/0406190`.

[18] V. E. Korepin, *Calculation of norms of Bethe wave functions*, Comm. Math. Phys. **86** (1982) 391–418.

[19] A. G. Izergin, V. E. Korepin, *The quantum inverse scattering method approach to correlation functions*, Comm. Math. Phys. **94** (1984), 67–92.

[20] A. G. Izergin, *Partition function of the six-vertex model in a finite volume*, Dokl. Akad. Nauk SSSR **297** (1987) 331–333; Sov. Phys. Dokl. **32** (1987) 878–879 (Engl. transl.).